Survey Options for Estimating Expenditure Weights for the Extramural Activities Component of the Biomedical Research and Development Price Index

FINAL REPORT

Submitted to:

Office of Science Policy
Office of the Director
National Institutes of Health

By:

Joel Popkin and Company
1155 15th Street, NW, Suite 614
Washington, DC 20005
(202) 872-0990
rkochhar@jpcecon.com

March 29, 2002

TABLE OF CONTENTS

Executive Summary

This report presents options for the estimation of expenditure weights for the extramural component of the BRDPI by means of a survey of institutions receiving funding from NIH. A survey is needed because institutions are no longer asked to submit data on research expenditures by budget categories during the grant application or renewal process for many types of awards. The research was accomplished via a review of the existing methodology, the statistical analysis of data from the Information for Management, Planning, Analysis, and Coordination (IMPAC) files, and the conduct of a test survey.

The analysis revealed that the unit of observation for the proposed survey should be the institution. Collection of data on the distribution of research expenditures on individual awards is not feasible. The expenditure survey can be counted upon to collect data on the same number of expenditure categories presently accounted for in the BRDPI, but no more. If desired, it is feasible to collect supplemental information in the survey, such as the proportion of salary and wages in an institution subject to the NIH salary cap. The questionnaire for the expenditure survey should adhere closely to the format developed for the test survey. The survey may be distributed by mail or e-mail. In-person interviews or visits to the institutions are not needed.

It is recommended that the expenditure weights in the BRDPI should be updated once every five years. The survey itself should be spread over a period of three to five years. At a minimum, the first one or two years of the survey cycle should focus on academic institutions and the next two years should be limited to non-academic institutions. An additional year might be needed to make up for non-response or take account of any other complications that might arise during the regular survey period. On average, each group of institutions should be surveyed once every five years. Spreading the survey period over three to five years will reduce the sample size in any single year and help eliminate the need for Office of Management and Budget (OMB) approval. It will also reduce the total burden of any single survey. Average respondent burden is expected to be in the range of five to nine hours.

A number of sampling strategies will provide excellent estimates of the expenditure weights, especially for academic institutions that account for almost 80 percent of total extramural funding. Because of the skewed nature of the distribution of NIH funding, a focus on institution size, as measured by total dollars awarded, is essential. For example, the nine largest academic institutions alone account for nearly 25 percent of total funding to academic institutions. Awards received by large institutions are also representative of the population of awards. A similar pattern is observed for non-academic institutions as well. Thus, small samples of institutions are sufficient to provide adequate coverage of extramural funding and accurate estimates of population expenditure shares. It is recommended that 9 to 18 large academic institutions and 9 to 27 large non-academic institutions be selected for the expenditure survey. If desired, a few smaller institutions might also be included in the sample without sacrificing much by way of accuracy. The survey sample in any single year need not exceed nine, thereby avoiding the need for OMB clearance. The small sample sizes will make it necessary for the Bureau of Economic Analysis (BEA) to make some revisions in the way it estimates the BRDPI. But none of those revisions are expected to compromise the accuracy of the BRDPI.

The estimated cost of a survey of the first nine institutions is approximately $9,000. This estimate includes start-up costs, such as, finalizing the survey design and drawing the sample. The cost of surveying 27 institutions over a period of three years is expected to be $18,500. The actual cost of the survey may fall outside of these boundaries depending on the tasks included in the final statement of work. For example, the cost estimates do not account for the statistical analysis, based on IMPAC data, of the characteristics of institutions and awards covered by and responding to the expenditure survey. On the other hand, some costs, such as the design of a survey, will be incurred infrequently and the cost of a second or third cycle of the survey may be lower than its initial cost.

DRAFT FINAL REPORT

Survey Options for Estimating Expenditure Weights for the Extramural Activities Component of the Biomedical Research and Development Price Index

1. Background and Major Issues

The Biomedical Research and Development Price Index (BRDPI) measures the change in prices of inputs purchased with the biomedical research budget of the National Institutes of Health (NIH). The BRDPI has two principal components: internal and extramural. The internal component covers intramural research and the administrative, planning and support activities performed by NIH in support of biomedical research. Extramural activities are performed by outside groups with support from the NIH in the form of research grants and awards. Such funding presently accounts for 80 percent of total NIH obligations for biomedical research. The outside groups are classified as academic or nonacademic. Academic institutions are primarily graduate and medical schools, while nonacademic organizations are diverse, ranging from private firms to nonprofit research organizations to State and local government agencies.

The BRDPI shares several traits with the Consumer Price Index (CPI) which is the most commonly used measure of inflation in the U.S. Both are input price indexes, i.e. they are weighted averages of the change in prices of a market basket of inputs. In the case of the CPI, the market basket consists of items bought by consumers for their personal use, such as, food, clothing, and transportation. The market basket for the BRDPI encompasses items needed for biomedical research, such as, research personnel, medical and laboratory equipment and supplies, and administrative support. Both indexes utilize weights that reflect the distribution of expenditures across the various items in their respective market baskets in the base period for the index. Such indexes are referred to as Laspeyres indexes. The current base period for the CPI is the 1993-95 time period and the BRDPI is based on FY 1993 expenditures. Finally, expenditure weights in both the CPI and the BRDPI are currently updated roughly every five years.[1]

[1] However, more frequent updating will soon be the norm in the CPI.

This report presents options for the estimation of expenditure weights for the extramural activities component of the BRDPI via the survey of institutions receiving extramural funding from NIH. Expenditure surveys would represent a new direction in the estimation of expenditure weights for the BRDPI. The existing methodology was developed by the Bureau of Economic Analysis (BEA) in 1983. The BEA methodology relied on the use of data from the Information for Management, Planning, Analysis, and Coordination (IMPAC) file maintained by NIH. Those data, which have traditionally been a rich source of information on extramural biomedical research expenditures by input categories, have become increasingly inadequate for the task. The reason is that information on the distribution of research expenditures by budget categories is no longer collected for non-competing awards. These awards currently represent over 50 percent of extramural expenditures on biomedical research at NIH, or about 40 percent of the total obligations for biomedical research. Moreover, the distribution of expenditures on research inputs is known to differ significantly depending upon whether an award was given on a competitive or non-competitive basis. The consequence is that the IMPAC data can no longer be used to determine a distribution of extramural biomedical research expenditures by input categories that is representative of all NIH obligations.

Because NIH no longer collects information on research expenditures by budget categories on non-competing awards, the data must be collected directly from the recipients of extramural awards by means of an expenditure survey. The survey will also be necessary in the future to make sure the weights do not become out of date. The major issues regarding the design of the survey are as follows:

(1) *Unit of observation*: In principle, the survey could collect data either on the distribution of expenditures on individual awards or the distribution of total funding received by individual institutions. There are several different types of awards distinguished by field of study, type of research (laboratory, clinical, animal, survey), duration of award, and the award mechanism (grant or contract, competing or non-competing, new or continuing, investigator initiated or program award.) There are also several different types of institutions distinguished by their type (academic or non-academic, private or public, hospital or research

institution), size, and location. The choice of the unit of observation depends partly on the source of the variance in the expenditure distribution of input categories. If differences across institutions do not matter and award characteristics are found to explain most of the variance, then the appropriate choice for the unit of observation is the award. However, practical considerations are also very important. It might be that data on individual awards are not readily available or that it is too burdensome to obtain those data. The institution then becomes a more suitable choice for the unit of observation.

(2) *Sampling strategy*: One of the major issues for the sampling strategy is whether the sample should be stratified. The appropriate choice of strata depends in part on the unit of observation and the distribution of expenditures across awards and institutions. Another issue is that of sample size. An unnecessarily large sample size will only add to respondent burden and survey costs.

(3) *Number of expenditure categories*: The number of expenditure categories used in the computation of the BRDPI at the present are constrained by the number of categories for which data are recorded in the IMPAC files. A survey offers the opportunity to add detail in this respect assuming the appropriate data are forthcoming from the institutions that are surveyed. It is, of course, desirable that the survey does not lead to a loss in the number of expenditure categories for which data are available.

(4) *Other data collection and respondent burden*: In addition to data on the expenditure distribution, a survey offers the opportunity to gather data on related issues that could affect the calculation of the BRDPI. One such issue is the salary cap imposed by NIH. This cap limits the extent of the upward drift in salary costs paid by NIH, a phenomenon not yet reflected in the BRDPI. However, nothing is known at this time about the resulting bias that might be present in the BRDPI. Another potential issue for research is the gathering of detailed data on indirect costs to sort price movements from quantity movements in changes in indirect

costs over time. The extent to which the survey could be used to address these issues depends on the burden it imposes on the respondent and the complexity that is added to the design of the questionnaire. Limiting the extent of the burden will be a key factor in increasing response rates. The total burden of the survey, an issue of interest to the Office of Management and Budget (OMB,) will also depend on the size of the sample.

(5) *Cost of the survey*: The cost of the survey depends upon the sample size, the survey method (personal visits, mail, e-mail or telephone,) and the length of the questionnaire and its level of complexity. A large sample size adds to the cost of developing the mailing list, printing and distributing the survey, and tabulating the responses. A more complex survey adds to the cost of editing and analyzing the data.

(6) *Preserving current methods of computing the BRDPI*: It is desirable that the expenditure survey is designed in a way that preserves continuity in the methods currently used by BEA to compute the BRDPI. Gathering data on the appropriate number of expenditure categories is an important part of meeting this objective. In addition, BEA's method for the academic component of the BRDPI requires institution-level detail on expenditures on wages, benefits, indirect costs and modified direct total costs (MDTC is the sum of expenditures on wages, benefits, consultants, travel and supplies.) These data are needed for over 150 of the largest institutions (ranked on the basis of funding received from NIH) included in BEA's sample. Of these data, indirect cost data are still available to BEA from the IMPAC files and other sources for all extramural institutions receiving NIH funding. The extent to which the remaining expenditure data will still be available to BEA depends on the final choice of sampling strategy. Depending on that choice, BEA may have to utilize imputation or otherwise modify its methods to adjust to the limits of the expenditure survey approach.

Research for this project was accomplished via a review of the existing methodology, the conduct of a test survey, and the analysis of data from IMPAC files for 1991, 1993 and 1995. The IMPAC data were used in part to test the efficacy of a range of sampling strategies using two principal criteria: the statistical variance of the sample estimates and the likely impact of using sample weights on estimated levels and rates of change in the BRDPI. In addition, alternative methods were tested for estimating indirect cost indexes at the institution level and for aggregating indirect cost, salary and wage, and fringe benefit indexes across institutions. Those are key components of the current BEA methodology but can only be supported by a fairly large sample of institutions. The purpose of testing alternative techniques was to ascertain whether a small sample of institutions would have an adverse effect on the BEA methodology and, by implication, on the BRDPI.

The major finding of the research is that the extramural BRDPI can be supported by a survey that collects data on the distribution of NIH funding across the same number of expenditure categories in use at the present time. Moreover, accurate data can be obtained by a survey of a relatively small number of institutions. That is feasible primarily because of the skewed nature of the distribution of NIH funding. The nine largest academic institutions alone account for nearly 25 percent of the total funding provided to academic institutions. Awards received by large institutions are also representative of the population of awards. Funding for non-academic institutions is also concentrated in the hands of a small proportion of institutions. As a result, a small sample of institutions is sufficient to provide adequate coverage of extramural funding and accurate estimates of population expenditure shares. In particular, a sample of 9 to 18 large academic institutions and 9 to 27 large non-academic institutions will be sufficient.

The appropriate unit of observation for the survey is the institution. Data on the distribution of expenditures on individual awards are either not forthcoming or too burdensome to collect. Several options are available for the selection of institutions to survey. Institutions may be stratified by size, as measured by total dollars awarded, and the strata with the large institutions could be over sampled. Non-probability methods, such as selecting the nine largest institutions, are also among the feasible alternatives. Finally, attention could be focused on the

100 largest academic institutions and the 100 largest non-academic institutions. Those institutions account for nearly all of the extramural funding provided by NIH. Small random samples from the "universe" of the top 100 institutions yield excellent approximations of the population weights. The sampling variance of the non-probability options is not defined but the other two leading options are associated with low variance and narrow confidence intervals for the sample weights. Simulations of the BRDPI using sample weights also lead to estimates very similar to those generated by BEA using population weights.

The small sample sizes will make it necessary for BEA to make some changes in the way it estimates the BRDPI. Those changes include the substitution of population data for institution data and using an alternative set of weights to aggregate indexes across institutions. But none of those revisions are expected to compromise the accuracy of the BRDPI. It is anticipated that OMB approval will not be required for the conduct of the expenditure survey. That is because the survey can be conducted over a period of three to five years with the sample size limited to no more than nine institutions in any single year. For example, a sample of 18 academic institutions could be covered over the first two years. Non-academic institutions could be surveyed over the succeeding years. On average, each group of institutions would be surveyed once every five years, which implies that BRDPI expenditure weights for the academic and non-academic components would also be updated with the same frequency. Small sample sizes also translate into modest costs. The cost of surveying the first set of nine institutions is expected to be $9,000 and the cost of surveying 27 institutions over a period of three years would be approximately $18,500. The small sample sizes also imply a limited amount of total respondent burden that is expected to be in the range of five to nine hours per respondent.

This report is organized into two major parts. The first part, consisting of seven sections, focuses on a broad overview of the research and the major findings of that research. The concluding section in this part contains recommendations regarding sample size, sampling strategies and the frequency and timing of the survey, estimates of the respondent burden, and projected costs associated with different sample sizes. The second major part of the report consists of a series of appendices that provide details on the research and its findings. The major topics covered by the appendices are as follows: (1) review of the current methodology; (2) the test survey; (3) analysis of the IMPAC data; (4) tests of alternative sampling strategies; (5) variance of sample estimates and BRDPI simulations with sample weights; and (6) the implications of the different sampling strategies for BEA's BRDPI methodology.

2. A Review of the Current Methodology

The extramural component of the BRDPI is estimated separately for academic and non-academic institutions. The academic index is comprised of nine expenditure categories and the non-academic index consists of eight expenditure categories. At the present time, price indexes for these categories are aggregated using weights derived from the 1993 IMPAC data. The institutions themselves were the source of the data as, prior to 1996, they were required to submit details on planned expenditures during the process of applying for NIH funding for any type of award. Table 1 below shows the existing expenditure categories and their weights for the extramural component of the BRDPI. Academic institutions (mainly institutes of higher education, such as, graduate and medical schools) account for approximately 80 percent of total NIH funding for extramural research.

Two points are apparent from a glance at Table 1. First, the category titled "other costs" has its weight set equal to zero by BEA even though expenditure data for that category was available in the IMPAC files. That is because it is not possible to design a corresponding price index for an "other cost" category and it is implicitly assumed that prices for this category move in the same fashion as a weighted average of all other direct cost categories. In reality, other costs comprise about 10 percent of total expenditures for academic institutions and about 15 percent of total expenditures for non-academic institutions. Second, indirect costs are excluded from the non-academic component of the BRDPI. Again, that is because there are no means by which BEA can estimate a suitable index for the change in indirect costs at non-academic institutions.

Table 1
Expenditure Categories and Their Weights in the Extramural Component of the BRDPI

Expenditure Category	*Academic Weight (a) 1993*	*Academic Weight (b) 1993*	*Non-Academic Weight 1993*
Total Costs	1.0000		
Total Indirect Costs	0.2843	0.0000	0.0000
Total Direct Costs	0.7157	1.0000	1.0000
Salary and Wages	0.4521	0.6317	0.6338
Fringe Benefits	0.1110	0.1551	0.1470
Consultant Services	0.0043	0.0060	0.0177
Equipment	0.0215	0.0300	0.0296
Supplies	0.1001	0.1399	0.1241
Travel	0.0133	0.0186	0.0233
Patient Care	0.0106	0.0148	0.0134
Alterations	0.0028	0.0039	0.0111
Other Costs	0.0000	0.0000	0.0000

Note: See the glossary attached to the report in Appendix I for definitions of the expenditure categories.
Source: BEA.

Table 1 also presents two sets of weights for academic institutions. The first set shows the weights as they are applied in the computation of the BRDPI. The second set shows the breakdown of direct costs within academic institutions for the sake of comparison with the expenditure distribution in non-academic institutions. The data show that there are scarcely any differences in the expenditure distributions at the two sets of institutions. Salary and wages, fringe benefits, and supplies are the dominant categories, accounting for over 90 percent of direct cost expenditures within both types of institutions. For academic institutions, indirect cost expenditures are also very important from the point of view of the BRDPI as they account for 28 percent of total expenditures, the second highest level after salary and wages.

Because of the dominance of the categories of salary and wages, fringe benefits, and indirect costs, BEA pays special attention to the construction of price indexes for them for the set of academic institutions. BEA collects data on salary and wages, fringe benefits, and indirect costs for each institution in a sample of 161 institutions that account for 95 percent of total disbursements to academic institutions. For the salary and wages component, BEA collects salary data for various ranks of faculty members classified further by whether or not they work in medical schools. Within institutions, salaries for the various types of faculty are weighted together using the numbers of each type that are employed in that institution. This is done separately for medical and non-medical schools. If an institution encompasses both types of schools, the average salaries at medical and non-medical school are weighted together based on the proportions of research dollars going to each type of school. Thus, BEA computes a salary and wages index for each of the 161 academic institutions in its sample. A weighted average of those indexes is then computed, with the weights based on the distribution of total salary and wages expenditures funded by NIH across the 161 institutions. The procedure for the fringe benefits index is similar in the sense that it also begins with the determination of fringe benefits at each institution. The fringe benefit indexes are then combined using weights derived from the distribution of total fringe benefit expenditures across the 161 institutions in the BEA sample.

Indirect cost indexes are computed for each academic institution as the product of two other indexes. One is an index reflecting changes in the indirect cost rate for that institution. The other is an index reflecting changes in the prices of direct costs on which overhead costs are paid. Those elements are referred to as Modified Direct Total Costs (MDTC) by NIH and consist of salary and wages, fringe benefits, supplies, consultants and travel. Thus, to compute an MDTC price index for each institution, BEA needs to know the distribution of expenditures across the direct cost components within each institution. Prior to 1996, that information could be readily obtained from the IMPAC files. Once an indirect cost index has been computed for each institution, they are weighted together based on the share of each institution in the total disbursements by NIH for indirect cost obligations. Data on disbursements of indirect cost obligations are still recorded in the IMPAC files.

The detailed procedures followed by BEA for computing price indexes for the salary and wages, fringe benefits, and indirect costs indexes require information on the expenditure distribution of NIH funding within each of the 161 academic institutions in the BEA sample.

Through 1995, this information was forthcoming from the IMPAC files. In principle, the same data may be gathered via an expenditure survey. Of course, a sample of 161 institutions would be considered very large given that the total number of academic institutions receiving NIH funding is only about 470. A smaller sample will clearly suffice because it is not imperative that the institutions that are selected account for as much as 95 percent of total extramural obligations to academic institutions. A more interesting question is whether alternative procedures are available for BEA so that it need not depend on data on the distribution of expenditures within individual academic institutions to compute indexes for indirect costs, salary and wages, and fringe benefits. It is shown below in Section 5 that this is indeed feasible. As a result, BEA's current procedures do not impose a constraint on the sample of academic institutions for the BRDPI expenditure survey. BEA's procedures for non-academic institutions are much simpler and do not require a breakdown of expenditures within individual institutions.

The remaining categories of expenditures – supplies, equipment, etc. – encompass a diverse array of items. For example, the price index for the equipment category is represented by a weighted average of the Producer Price Indexes (PPI) for technical, scientific and professional books, commercial furniture, non-wood furniture, and surgical and medical instruments and apparatus. The weights are based on data from the IMPAC file, interviews with NIH staff, and other miscellaneous sources of data. The price index for the supplies category is even more internally heterogeneous than the index for equipment.[2] As a result, the estimation of the BRDPI could benefit from an expansion of the number of categories for which expenditure weights are estimated.[3] In particular, it would be valuable to add a category for computer equipment and services. That is because the price behavior of these items is known to differ sharply from that of the remaining items. Thus, the expansion of expenditure categories in the extramural component of the BRDPI is a worthwhile objective for the expenditure survey and its feasibility was tested with the test survey. Unfortunately, the test survey showed that the expenditure survey can be counted upon to collect data on the same number of expenditure categories presently accounted for in the BRDPI, but no more (see Section 3 below for details.)

[2] See Appendix A for details.

[3] An outside target for this expansion is to aim for the same level of detail that is used for the internal activities component of the BRDPI. This component of the BRDPI consists of 27 expenditure categories. Those categories and their weights are shown in Appendix A.

Another issue that can be explored by an expenditure survey is the effect of the salary cap on the BRDPI. The salary cap refers to the maximum annual salary rate, exclusive of fringe benefits and facilities and administration costs, that the NIH will pay for any single individual under an NIH award.[4] The principal implication of the salary cap for the BRDPI is that, as long as it is binding for some research staff, it limits the growth in salary and wages faced by NIH in comparison to the actual growth in salary and wages that is currently captured by the BRDPI. Some idea of the potential bias in the BRDPI salary and wages component may be gleaned from the survey by making broad inquiries of respondents, such as, the percentage of salary and wages at the institution that are subject to the salary cap.[5] The test survey showed that it is feasible to gather data on the effect of the salary cap from institutions but it is a potentially burdensome line of inquiry.

3. The Test Survey

The test survey was a major step towards the development of a final strategy for the proposed BRDPI expenditure survey. The primary objectives of the test survey were as follows: (1) Determine design, content and feasibility of the final expenditure survey. (2) Estimate the likely burden on survey respondents. (3) Assess the feasibility of extending the number of expenditure categories beyond those currently in use. (4) Field-test the definitions of expenditure categories and other items for which data are to be collected. (5) Test the feasibility of using either the award or the institution as the unit of observation. (6) Determine the feasibility of extending the scope of the expenditure survey to cover related issues of interest, such as, the impact of the salary cap on limiting the growth in wage costs for extramural research. (7)

[4] The salary cap is not constant and changes over time. In recent times the salary cap has increased from $125,900 (Federal FY 1999 awards approved before 12/31/99) to $161,200 (Federal FY 2001 awards approved after 1/1/01). The salary cap assumes a full-time, 12-month workload. Thus, to determine if the salary cap might apply to any individual, his or her institutional base salary for the time period in question must be converted to a full-time, 12-month basis. For example, the salary of a person on a half-time appointment must be doubled to determine whether the salary cap is binding. Also, the limitation is prorated to the level of effort (percentage of time) requested for the research project. For example, the maximum amount of salary that could be charged in FY 1999 for a person on a full-time, 12-month appointment devoting 10 percent of his or her time to the NIH grant is $12,590.

[5] As shown above by the outline of the functioning of the salary cap, it impacts the salaries institutions may charge to NIH grants in a fairly complex manner and it is unlikely that an expenditure survey could reasonably uncover the full extent of the impact on each institution.

Determine the survey technique – mail, e-mail, phone, or personal visits – that is preferred by respondents and, therefore, is most likely to succeed.

The major accomplishment of the test survey is that it proved the feasibility of using the survey as an instrument for collecting data on the expenditure distribution of extramural research funding. A key difference between the survey and the present reliance on IMPAC data is that the latter yields the distribution of *planned* expenditures whereas the former will estimate the distribution of *actual* expenditures. Planned expenditures, which correspond to funds obligated by NIH in a given Federal fiscal year, are not necessarily incurred in that year or even within a single fiscal year. Thus, respondents would have to trace expenditures for individual awards over time, a process that would be especially burdensome on the institutions receiving several hundred awards a year from NIH. Past research conducted by NIH has shown that the distributions of planned and actual expenditures are not identical. Thus, there will be some impact on the level and rate of change in the BRDPI as survey weights replace IMPAC-based weights. But that impact is expected to be very small.

The test survey also demonstrated the feasibility of collecting data for the same number of expenditure categories that are contained in IMPAC. Additional detail is either not available or would prove too burdensome to provide. If desired, an expenditure survey can also be used to obtain information on related issues such as the salary cap. However, those questions do add to the burden of the survey. It is estimated that the BRDPI expenditure survey will take the average respondent nine hours to complete. The response time will shrink considerably if the survey is confined strictly to questions on the distribution of expenditures.

A key issue that was settled by the test survey is that it is not feasible to use the award as the unit of observation. This might have been desired, for example, if award characteristics account for much of the variance in the distribution of expenditures, an issue taken up in Section 4. But even if that is the case, the test survey shows that the unit of observation in future surveys must be the institution.

The remainder of the discussion in this section focuses on the key aspects of the test survey. Further details are presented in Appendix B. The test survey and its attachments – a cover letter and a glossary – are reproduced in Appendix I. The initial version of the test survey requested expenditure detail on a fairly detailed list of categories (see Appendix B.) The early version also retained both the institution and the award as the unit of observation. Thus, this version of the test survey requested the distribution of total expenditures at an institution, and the distribution of expenditures on each award received by an institution during the reference year for the survey. The draft version of the test survey was then circulated for comment within NIH, two large academic institutions, the Council of Governmental Relations (COGR), and the Association of American Medical Colleges (AAMC.)

Comments on the draft version of the test survey indicated strongly that data on the distribution of expenditures on individual awards are either not available or, if available, would be very burdensome to produce. Requesting data for just a sample of awards might, if anything, complicate the task for the respondents because they would have to search and extract data from within a larger accounting system. The unanimity of the viewpoints on this issue easily settled the choice between the institution and the award as the unit of observation in favor of the institution. There was also an overwhelming negative reaction on the part of potential respondents with respect to the level of expenditure detail on which data were requested in the draft test survey. The typical reaction was that either the detail is not available or that record keeping at individual institutions will differ in varying fashions from the categories listed in the draft version. Therefore, it was decided to simplify the information request on the expenditure distribution. Commentators also had a number of reactions to the definitions of expenditure terms and concepts. Those were reviewed carefully and revisions were made accordingly.

The final test survey was mailed to nine institutions in July 2001. That is the largest sample size possible without triggering the need to obtain OMB clearance. The selected institutions were regionally diverse, with two each from the northeast, mid-Atlantic, west, and mid-west regions, and one from the south region. The nine institutions also represented institutes of higher education, hospitals, private research institutions, and government organizations.

Within the selected regions and types of organizations, the nine institutions are among those receiving a large share of awards and funding.

The list of expenditure categories in the final version of the test survey is the same as that presently in use in the BRDPI with one exception: the test survey has a line item for consortium/contractual costs.[6] There is also a line item for fee/profit that applies to contracts only. The presence or absence of this item is not expected to have a significant impact on estimates of expenditure weights because fees/profits make up only a small percentage of obligations for contracts which, in turn, account for less than 10 percent of total NIH obligations.

The test survey also included a question on the percentage of salaries and wages paid to personnel subject to the NIH salary cap. If it is assumed that the mean answer to this question is 20 percent, that share of the expenditure weight for wages and salaries could be assumed to correspond to zero percent growth in wages in between changes in the salary cap. That will yield an estimate of the short-run upward bias in the BRDPI arising from the inability to capture the effect of the salary cap. The salary cap does move upwards in the long run, albeit at discreet intervals. Thus, the extent of the bias in the long run depends on the average increase in the salary cap over time relative to the overall average growth in wages. Other key questions in the test survey focused on issues related to estimating the burden of the survey.

The test survey generated responses from seven out of the nine institutions. According to IMPAC data for 1999, these seven institutions collectively received $1.09 billion in funding, or nearly 10 percent of the total of $11.30 billion in extramural funds obligated by NIH in Federal FY 1999. Out of the total of 39,820 awards approved for funding, 3,441 (or 8.6 percent) were issued to these seven institutions. Thus, these seven institutions account for a sizable share of funding and awards, and the expenditure shares estimated from the test survey could be treated as a preliminary indicator of the results that might emerge from a full survey. The major findings from the test survey are summarized in Table 2. More detailed results are shown in Appendix B.

[6] See the glossary in Appendix I for the definition of consortium/contractual costs.

Table 2
Expenditure Shares and Related Data from the
Test Version of the BRDPI Expenditure Survey, 2001

Expenditure Shares	*Survey Averages* Institutions 1 to 6	Institutions 1 to 7	*Shares from IMPAC 1995* Institutions 1 to 6	Institutions 1 to 7	All Institutions
Total Costs	1.0000	1.0000	1.0000	1.0000	1.0000
Total Indirect Costs	0.2937	0.3055	0.3357	0.3411	0.2975
Total Compensation	0.3965	0.4007	0.4434	0.4403	0.4592
Salary and Wages	0.3216	---	0.3506	0.3501	0.3702
Fringe Benefits	0.0749	---	0.0928	0.0902	0.0890
Consultant Costs	0.0030	0.0027	0.0029	0.0028	0.0050
Equipment	0.0247	0.0243	0.0150	0.0151	0.0174
Supplies	0.0974	0.0953	0.0893	0.0905	0.0819
Travel Costs	0.0112	0.0105	0.0087	0.0085	0.0119
Patient care	0.0147	0.0140	0.0094	0.0089	0.0084
Alterations and renovations	0.0036	0.0032	0.0004	0.0004	0.0007
Other costs	0.1553	0.1439	0.0950	0.0925	0.1180
Fee/profit	0.0000	0.0000	---	---	---
Percentage of salary and wages Subject to NIH salary cap		2.6 – 13.1%			
Hours to complete survey					
Test survey	13	23			
Similar survey in future	9	19			

Note: Other costs include consortium/contractual and miscellaneous expenses.
Source: Joel Popkin and Company test expenditure survey of nine institutions.

In the interests of making a full comparison of the results from the test survey with data available from the IMPAC files, the expenditure distribution in Table 2 retains the expenditure category of other costs. Six out of the seven institutions were able to report all of the requested data. The seventh was unable to provide a breakdown of compensation into wages and salaries and fringe benefits. A follow-up query with this institution confirmed that this was not simply an omission on their part and reflected a genuine inability to provide the data. Thus, the expenditure shares data in Table 2 are presented both with and without the inclusion of this institution.

For the seven institutions combined, the expenditure share of indirect costs is 30 percent, the total compensation share is 40 percent, the share of other costs is 14 percent, and supplies account for nearly 10 percent.[7] These data, pertaining to FY 2000, can be compared to expenditure shares computed from the IMPAC file for 1995, the last year that detailed expenditure data were recorded in the IMPAC files. It can be seen from Table 2 that the survey data show lower expenditure shares for indirect costs and compensation, and higher shares for other costs. At least part of this difference is the five-year gap between the date of the IMPAC file and the date of the survey. However, the majority of the difference, no doubt, is due to the gap between actual and planned expenditures. Past research by NIH has shown that, in contrast to planned expenditures, actual expenditures are distributed away from labor compensation and towards other components, such as, equipment and supplies. The pattern in Table 2 shows that there is also redistribution towards indirect costs (an element not covered by the NIH research) and towards other costs. The last column in Table 2 shows expenditure weights computed for the entire population of institutions from the IMPAC data for 1995. Apart from the differences just noted, there is fairly good correspondence between the weights estimated from the test survey and the weights for the extramural population. This provides good reason to be optimistic about the prospects for acquiring reliable estimates of expenditure weights from a full-scale expenditure survey.

The test survey also asked about the percentage of salary and wages that was subject to the NIH salary cap. As shown in Table 2, this ranged from a low of 2.6 percent to a high of 13.1 percent. It is impossible to say whether this range is typical of the population of institutions, but these data do suggest that the inability to account for the salary cap in the BRDPI at the present time may not lead to a significant upward bias.

An important subject for the test survey was to determine the burden on the respondent. Measured in terms of the hours taken to complete the survey, the burden for the test survey ranged from 2 to 80 hours. The top end of the range was surprisingly high and an inquiry was

[7] The weights are "plutocratic" weights. In other words, institutions receive a weight in proportion to the total dollars awarded to them. That is equivalent to summing the expenditures on a given category, say, supplies, across all institutions and then dividing by the total dollars awarded to all institutions to estimate the weight for that

made with this institution to verify the reported number. It was determined that the reported number is accurate and is that high because accounting records at that institution are not currently computerized in a manner suitable for completing the BRDPI expenditure survey. With reprogramming, which the institution would undertake if it had to regularly complete the survey in the future, the burden would be much less. Therefore, the burden reported by this institution is best treated as an outlier. Excluding this institution, the average burden for the completion of the test survey was 13 hours. Institutions were also asked if a future survey might take less time to complete since they had now become familiar with the concepts being addressed by the survey and had done the programming necessary to accomplish the task. Four out of the six institutions reported that a future survey would take less time to complete, and the estimated average time to complete this survey in the future is nine hours (excluding the outlier of 80 hours.) Some institutions commented that the question on the salary cap consumed much of the time it took to complete the survey. Therefore, if that question were to be omitted, one can expect the burden in the future, at least for the present group of six institutions, to be below nine hours, possibly as little as five hours.

4. The Analysis of IMPAC Data

This phase of the research consisted of the analysis of IMPAC data for 1991, 1993 and 1995.[8] The primary objective was to tabulate population expenditure weights and determine their relationship to the characteristics of institutions and awards. This relationship is worth examining for the development of a sampling frame and a stratification plan for the expenditure survey. Three different years were chosen for this purpose to examine whether these relationships were stable over time. The analysis found that the properties of the expenditure weights held steady for the duration of the 1991 to 1995 time period. Therefore, data sometimes are presented only for 1993. While that is not the latest year for which IMPAC data with expenditure detail are

category. That is also the procedure followed by BEA and throughout this paper unless explicitly mentioned otherwise.

[8] IMPAC data for 1999 were also acquired from NIH. The data for 1999 are of limited use because they do not contain detailed expenditure information, but they were used to examine the current distribution of awards and institutions by their major characteristics. Because no changes of note were found to have occurred between 1995 and 1999, the discussion in this section is confined to the findings from the data through 1995.

available, it is the year that corresponds to the date for the weights in the BRDPI at the moment. Additional data are presented in Appendix C.

The distribution of expenditure weights was examined through a variety of means, including the analysis of variance (ANOVA) and regression analysis. The output from the analyses of the IMPAC data from the various time periods is voluminous and is not discussed or reproduced here in its entirety. Instead, the focus of this section is on summarizing the principal findings and discussing their implications for the expenditure survey. A leading issue for analysis was whether the characteristics of institutions and awards play a role in shaping the distribution of expenditure weights. Determining which of the two – institutions or awards – is more important and which characteristics were most significant would help settle the choice of the unit of observation and assist in the design of a stratified sample. However, the discussion of the ANOVA and regression analysis that was conducted for this purpose is mostly confined to Appendix F for two reasons. One reason is that the statistical analysis was inconclusive in the sense that most characteristics of awards and institutions were found to play some role in the determination of expenditure weights. Thus, no clear-cut short list of institution or award characteristics emerged as potential stratifying variables. The second reason is that the choice of the unit of observation was settled in favor of the institution by the test survey. Thus, the statistical analysis proved to be of less import than originally thought.

The most important finding of the analysis of IMPAC data is that the size of an institution, as determined by the dollars in funding received by an institution from NIH, is a critical factor for the design of the BRDPI expenditure survey. It is found that both awards and dollars are concentrated in the hands of relatively few academic and non-academic institutions and in relatively few types of awards. From the point of view of estimating expenditure weights, the key consideration is not the proportion of institutions or awards that is captured by a sample but the share of dollars represented by that sample. Thus, the size of an institution – as measured by dollars granted or awards received – needs to be an important consideration in the design of a sample for the expenditure survey. Because of the skewed nature of the distribution of NIH funding, a survey of a handful of institutions has the potential of yielding very accurate results. That is particularly true of samples stratified by size of institution, a point covered in Section 6.

4.1. The Distribution of Extramural Funding

Table 3 below shows the number of institutions and the total number of awards and dollars awarded to those institutions in 1991, 1993 and 1995.[9] The data show that while academic institutions comprise only 30 percent of all institutions receiving extramural funding from the NIH, they receive about 80 percent of the awards and total dollars awarded. The shares of non-academic institutions have been on the increase since 1991, but at a slow pace. There is also a sharp disparity between academic and non-academic institutions in the number of awards received per institution. On average, an academic institution receives slightly over 50 awards whereas the typical non-academic institution receives only 5 awards. Because of the dominant share of academic institutions in both dollars and awards, the implication of Table 3 is that greater effort should be focused on measuring the expenditure weights for academic institutions than non-academic institutions. Alternatively, the estimation strategy could tolerate a greater degree of statistical error in the measurement of expenditure weights for non-academic institutions.

The distribution of academic and non-academic institutions by dollars awarded and number of awards received is shown in Table 4 (see Appendix D for additional detail.) Within academic institutions, the distribution of dollars is highly skewed towards the top end. In 1993, 33 out of 472 academic institutions in 1993 received over $50m. in funding each and accounted for 57 percent of total dollars received by academic institutions. Indeed, the 107 institutions (or 23 percent of all academic institutions) that each received over $10m. in funding accounted for over 90 percent of the total funding to academic institutions. The distribution of non-academic institutions by dollars awarded is also quite skewed. In 1993, 5 out of 1,150 non-academic institutions controlled over 21 percent of the total dollars awarded. There were 28 institutions (or less than 3 percent of the total) receiving over $10m. in funding each and they collectively accounted for 51 percent of total funding given to non-academic institutions. Nearly three-

[9] It should be noted that these tabulations were derived from IMPAC files that had been edited with the intent of estimating expenditure weights for the extramural portion of the BRDPI. Thus, awards of certain activity types or observations with missing data on expenditures were deleted from the files. Similarly, institutions located abroad were eliminated from consideration. The totals shown in Table 3, therefore, do not represent the universe of institutions receiving extramural funding from the NIH or the total number of awards and dollars received by them.

Table 3
Number of Institutions, Number of Awards, Total Dollars Awarded and Their Percentage Distribution

	1991		1993		1995	
Number of Institutions						
Academic	473	30.4%	472	29.1%	474	27.6%
Non-Academic	1,081	69.6%	1,150	70.9%	1,239	72.3%
Total	1,554	100.0%	1,622	100.0%	1,713	100.0%
Number of Awards						
Academic	24,005	80.6%	24,400	80.0%	24,701	79.4%
Non-Academic	5,776	19.4%	6,101	20.0%	6,414	20.6%
Total	29,781	100.0%	30,501	100.0%	31,115	100.0%
Total Dollars Awarded						
Academic	$5,156,227,809	79.2%	$5,745,461,678	78.6%	$6,222,049,487	78.1%
Non-Academic	$1,356,086,241	20.8%	$1,562,344,361	21.4%	$1,741,437,738	21.9%
Total	$6,512,314,050	100.0%	$7,307,806,039	100.0%	$7,963,487,225	100.0%

Source: Joel Popkin and Company tabulations from IMPAC data.

quarter of non-academic institutions received less than $0.35m. each in funding and accounted for less than 9 percent of the total funding.

The patterns that are evident with respect to the distribution of dollars awarded are mirrored in the distribution of awards. That, of course, is not surprising since one would expect total funding received to be closely related to the total number of awards received. In 1993, 78 academic institutions (or 16 percent of all academic institutions) received over 100 awards each. These institutions also accounted for 83 percent of total funding. As previously noted, non-academic institutions generally receive a fewer number of awards with over one-half receiving only one award each. Roughly 10 percent of non-academic institutions receive 10 or more awards each and they account for approximately 75 percent of the dollars awarded. The data in Table 4 are the first indication that expenditure weights for both types of institutions could be reliably estimated by focusing on a handful of large recipients.

Table 4
Distribution of Total Dollars Awarded Across Institutions Grouped by Dollars Awarded and Number of Awards Received, 1993

	Academic Institutions			*Non-Academic Institutions*		
Dollars Awarded	Number of Institutions	Total Dollars Awarded	Percent Distribution	Number of Institutions	Total Dollars Awarded	Percent Distribution
$50 m. or more	33	$3,280,896,882	57.10	5	$335,423,796	21.47
$25 to 49 m.	32	$1,221,563,283	21.26	4	$159,190,369	10.19
$10 to 24 m.	42	$710,202,482	12.36	19	$299,336,641	19.16
$5 to 9 m.	38	$260,565,485	4.54	36	$247,313,112	15.83
$2.5 to 4 m.	33	$111,566,614	1.94	53	$186,173,551	11.92
$1 to 2.4 m.	54	$82,612,669	1.44	69	$110,113,589	7.05
$0.75 to 0.9 m.	20	$17,154,705	0.30	44	$37,692,769	2.41
$0.5 to 0.74 m.	36	$21,804,816	0.38	81	$49,494,120	3.17
$0.35 to 0.49 m.	28	$12,322,008	0.21	79	$33,497,596	2.14
$0.25 to 0.34 m.	38	$11,245,082	0.20	128	$37,540,899	2.40
$0.1 to 0.24 m.	74	$12,169,407	0.21	260	$46,386,126	2.97
Less than $0.1 m.	44	$3,358,245	0.06	372	$20,181,793	1.29
Total	472	$5,745,461,678	100.00	1150	$1,562,344,361	100.00

	Academic Institutions			*Non-Academic Institutions*		
Number of Awards	Number of Institutions	Total Dollars Awarded	Percent Distribution	Number of Institutions	Total Dollars Awarded	Percent Distribution
250 or more	27	$2,921,854,353	50.85	3	$233,497,883	14.95
100 to 249	51	$1,852,399,327	32.24	6	$261,116,282	16.71
50 to 99	32	$453,562,987	7.89	11	$186,661,000	11.95
25 to 49	40	$255,025,950	4.44	31	$238,270,003	15.25
10 to 24	49	$129,205,620	2.25	74	$265,915,170	17.02
5 to 9	65	$69,589,156	1.21	84	$112,209,749	7.18
2 to 4	91	$43,500,184	0.76	313	$155,765,671	9.97
1	117	$20,324,101	0.35	628	$108,908,603	6.97
Total	472	$5,745,461,678	100.00	1150	$1,562,344,361	100.00

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 5
Percentage of Dollars and Awards Received by the Top 9, Top 20 and Top 100 Institutions

	Percentage of Dollars Awarded to:			*Percentage of Number of Awards to:*		
Type of Institution	Top 9 Institutions	Top 20 Institutions	Top 100 Institutions	Top 9 Institutions	Top 20 Institutions	Top 100 Institutions
Academic						
1991	23.5	42.3	89.1	19.6	36.0	89.0
1993	23.7	42.7	89.4	20.1	37.0	88.1
1995	23.7	43.0	89.5	20.9	38.1	89.5
Non-Academic						
1991	30.9	44.7	75.3	25.1	36.4	69.8
1993	31.7	44.7	75.6	27.3	37.9	68.3
1995	31.1	44.7	76.1	28.8	40.3	70.1
All Institutions						
1991	18.6	33.7	86.2	15.8	28.9	85.3
1993	18.7	33.8	86.4	16.1	29.6	84.2
1995	18.5	33.9	86.6	16.6	30.4	85.5

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 5 provides an alternative perspective on the concentration of funding and awards in the hands of the top institutions ranked by dollars awarded. The table shows the percentage of total dollars awarded going to the top nine, top 20 and top 100 institutions within that category. A similar breakdown is shown for the number of awards approved. The reason for this presentation is that these groups of institutions comprise potential samples of institutions for the expenditure survey. The top nine institutions are of particular interest because a sample of nine or fewer institutions does not require OMB clearance.

The data in Table 5 show that just the top nine or top 20 institutions receive a very significant share of total dollars awarded to both academic and non-academic institutions. The same is true for the number of awards received by these institutions. Thus, it is possible that reliable estimates of the population expenditure weights could be obtained by sampling either just the top nine or top 20 institutions within the set of academic and non-academic institutions. The group of the top 100 institutions controls almost all the dollars and awards issued by NIH and virtually forms an alternative "universe" of institutions. In other words, a sampling strategy designed to estimate expenditure weights for the top 100 institutions is potentially the same as attempting to estimate the weights for the entire population of institutions. Sampling experiments with these select groups of institutions are presented in Section 6 of the report.

Appendix D contains additional tabulations on the distribution of institutions and awards by their major characteristics and the flow of funding to each type of institution or award. The overall picture that emerges from those tabulations is one of concentration. In addition to the overwhelming importance of size, both awards and dollars are concentrated in the hands of relatively few types of academic and non-academic institutions (e.g. institutions of higher education or non-profit, private independent organizations) and relatively few types of awards (e.g. non-competing continuations or Type R awards.) From the point of view of estimating expenditure weights, the key consideration is not the proportion of institutions or awards that is captured by a sample but the share of dollars represented by that sample. Thus, the size of an institution – as measured by dollars granted or awards received – needs to be an important consideration in the design of a sample for the expenditure survey. Especially within non-academic institutions, it is apparent that a simple random sample would draw primarily from small institutions and individual recipients of NIH funding that tend to dominate in numbers but are insignificant with respect to dollars awarded. The characteristics of awards and institutions also did not vary much between 1991 and 1995. While not explicitly noted in the tables presented here, this stability also extends through 1999. Thus, a sampling frame derived from the IMPAC data need not be revisited too often.

4.2. Expenditure Weights and Their Distribution

The expenditure weights computed and used by BEA for the estimation of the BRDPI were shown above in Table 1. This section presents population expenditure weights computed for this research from the IMPAC files and shows how those weights vary across institutions.

Some of the data in this section are for 1993 only. Data for 1991 and 1995 are very similar and may be found in Appendix C. That appendix also contains detailed tabulations on expenditure weights computed by major characteristics of institutions and awards.

Like the BEA, all expenditure weights were computed using the "plutocratic" method, i.e. by giving a larger weight to institutions receiving a higher amount of funding from NIH. There are two (equivalent) ways of achieving this result. Consider the example of estimating the wage expenditure share. One method would be to sum the wage expenditures of all institutions and dividing that by the sum of total expenditures in all institutions. The other method would be to first compute the wage expenditure share within each institution. A weighted average of the wage shares can now be taken where the weights reflect the proportion of total expenditures received by each institution. The plutocratic method was applied uniformly throughout the research except in the case of the analysis presented in Appendix E. For reasons given in that appendix, the shares used in that analysis are "democratic" weights, i.e. they represent unweighted averages of shares across institutions.

Table 6 contrasts the estimated population expenditure weights with those used by BEA for the BRDPI. The weights are expressed so that the direct cost components sum to one. As is the case in the extramural BRDPI, the weight for other costs is set equal to zero. An additional adjustment is made for non-academic institutions by removing the indirect cost category. Table 6 shows that the population weights computed from the 1991, 1993 and 1995 IMPAC files are very similar to the weights originally estimated by BEA for use in the BRDPI. That is the case for both academic and non-academic institutions. The weights from the 1993 IMPAC files are not identical to the BEA weights because of slight differences in the IMPAC files provided to us by NIH and those provided to BEA. In particular, the number of observations reported by BEA are different from the number in the IMPAC files we received. The editing procedures followed by BEA are also slightly different.

Tables 7(a) and 7(b) present the major characteristics of the distributions of expenditure shares using the institution as the unit of observation. In other words, the statistics in the tables – percentiles, standard deviations, etc. – are computed across institutions. For both academic and

Table 6
Expenditure Weights for Academic and Non-Academic Institutions Compared With BEA Weights

Expenditure Category	*Academic Institutions*			
	1991	1993	1995	BEA: 1993
Indirect costs	0.3000	0.2967	0.2962	0.2843
Direct costs	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6337	0.6362	0.6334	0.6317
Fringe benefits	0.1415	0.1554	0.1537	0.1551
Consultants	0.0058	0.0061	0.0065	0.0060
Equipment	0.0460	0.0284	0.0299	0.0300
Supplies	0.1387	0.1401	0.1420	0.1399
Travel	0.0179	0.0184	0.0191	0.0186
Patient care	0.0141	0.0137	0.0140	0.0148
Alterations, etc.	0.0023	0.0017	0.0013	0.0039
Other	0.0000	0.0000	0.0000	0.0000
	Non-Academic Institutions			
	1991	1993	1995	BEA: 1993
Indirect costs	0.0000	0.0000	0.0000	0.0000
Direct costs	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6486	0.6423	0.6330	0.6338
Fringe benefits	0.1420	0.1483	0.1468	0.1470
Consultants	0.0144	0.0176	0.0169	0.0177
Equipment	0.0342	0.0291	0.0290	0.0296
Supplies	0.1213	0.1250	0.1331	0.1241
Travel	0.0205	0.0232	0.0251	0.0233
Patient care	0.0184	0.0139	0.0158	0.0134
Alterations, etc.	0.0007	0.0005	0.0004	0.0111
Other	0.0000	0.0000	0.0000	0.0000

Source: BEA and Joel Popkin and Company tabulations from IMPAC data.

Table 7(a)
Major Characteristics of the Distributions of Expenditure Shares:
Academic Institutions, 1993

Expenditure Category	*Mean*	*Minimum*	*Maximum*	*Percentiles* 25th	*Percentiles* 50th	*Percentiles* 75th	*Standard Deviation*	*Coefficient of Variation*
All Institutions								
Indirect Costs	0.2967	0.0000	0.4855	0.2395	0.2749	0.3027	0.0917	0.3091
Direct Costs								
Salary and Wages	0.6362	0.0000	1.0000	0.5970	0.6418	0.6777	0.0665	0.1046
Fringe Benefits	0.1554	0.0000	0.2730	0.1118	0.1412	0.1629	0.0532	0.3422
Consultants	0.0061	0.0000	0.2330	0.0000	0.0047	0.0135	0.0047	0.7758
Equipment	0.0284	0.0000	0.6008	0.0125	0.0294	0.0610	0.0248	0.8722
Supplies	0.1401	0.0000	0.3933	0.0913	0.1344	0.1613	0.0452	0.3223
Travel	0.0184	0.0000	1.0000	0.0161	0.0206	0.0300	0.0125	0.6801
Patient Care	0.0137	0.0000	0.0962	0.0000	0.0000	0.0007	0.0224	1.6376
Alterations, etc.	0.0017	0.0000	0.0940	0.0000	0.0000	0.0000	0.0119	6.9798
Top 100 Institutions								
Indirect Costs	0.2989	0.2003	0.3673	0.2706	0.2965	0.3218	0.0469	0.1571
Direct Costs								
Salary and Wages	0.6360	0.5488	0.6959	0.6178	0.6355	0.6580	0.0341	0.0536
Fringe Benefits	0.1569	0.0908	0.2301	0.1366	0.1565	0.1682	0.0272	0.1735
Consultants	0.0056	0.0011	0.0134	0.0038	0.0055	0.0072	0.0021	0.3820
Equipment	0.0265	0.0074	0.0804	0.0201	0.0245	0.0363	0.0117	0.4411
Supplies	0.1404	0.0695	0.2228	0.1279	0.1414	0.1546	0.0230	0.1640
Travel	0.0179	0.0103	0.0429	0.0153	0.0174	0.0202	0.0064	0.3567
Patient Care	0.0149	0.0000	0.0633	0.0011	0.0112	0.0195	0.0113	0.7582
Alterations, etc.	0.0018	0.0000	0.0376	0.0000	0.0000	0.0002	0.0061	3.4082

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 7(b)
Major Characteristics of the Distributions of Expenditure Shares:
Non-Academic Institutions, 1993

Expenditure Category	*Mean*	*Minimum*	*Maximum*	*Percentiles*			*Standard*	*Coefficient*
				25th	*50th*	*75th*	*Deviation*	*of Variation*
All Institutions								
Indirect Costs	0.3102	0.0000	0.6867	0.0289	0.2098	0.3374	0.2675	0.8624
Direct Costs								
Salary and Wages	0.6423	0.0000	1.0000	0.5650	0.6550	0.7290	0.1836	0.2859
Fringe Benefits	0.1483	0.0000	0.5034	0.0528	0.1276	0.1673	0.1233	0.8316
Consultants	0.0176	0.0000	0.7272	0.0000	0.0093	0.0650	0.0690	3.9225
Equipment	0.0291	0.0000	1.0000	0.0000	0.0068	0.0584	0.0714	2.4532
Supplies	0.1250	0.0000	0.7264	0.0164	0.0654	0.1556	0.2011	1.6085
Travel	0.0232	0.0000	1.0000	0.0000	0.0143	0.0333	0.0669	2.8855
Patient Care	0.0139	0.0000	0.8322	0.0000	0.0000	0.0000	0.0720	5.1780
Alterations, etc.	0.0005	0.0000	0.1183	0.0000	0.0000	0.0000	0.0162	32.3508
Top 100 Institutions								
Indirect Costs	0.3283	0.0177	0.4909	0.2599	0.3170	0.3785	0.0877	0.2671
Direct Costs								
Salary and Wages	0.6428	0.2415	0.8394	0.6122	0.6480	0.6768	0.0698	0.1086
Fringe Benefits	0.1534	0.0000	0.3248	0.1306	0.1591	0.1845	0.0475	0.3097
Consultants	0.0106	0.0000	0.3977	0.0038	0.0070	0.0182	0.0182	1.7154
Equipment	0.0222	0.0000	0.2599	0.0075	0.0166	0.0268	0.0229	1.0293
Supplies	0.1361	0.0000	0.4411	0.0725	0.1178	0.1681	0.0729	0.5359
Travel	0.0194	0.0052	0.2586	0.0117	0.0151	0.0245	0.0205	1.0566
Patient Care	0.0148	0.0000	0.2191	0.0000	0.0000	0.0096	0.0266	1.7944
Alterations, etc.	0.0007	0.0000	0.0335	0.0000	0.0000	0.0000	0.0064	9.1998

Source: Joel Popkin and Company tabulations from IMPAC data.

non-academic institutions, Tables 7(a) and 7(b) show estimates for the population of institutions and the top 100 institutions based on dollars awarded. For the sake of brevity, estimates for 1991 and 1995 are presented in Appendix C.

Looking at the data for all institutions in Table 7(a), it is apparent that there is fair amount of variation in the expenditure share for any category. The standard deviations are often on the high side, as are the coefficients of variation (CV.)[10] A striking feature of the data in Table 7(a) is the virtual identity between the expenditure weights for all institutions and for the top 100. That is not surprising, however, because it was noted earlier that the top 100 academic institutions receive 89 percent of the total funding given to all academic institutions. Even more striking is the fact that the variance of the expenditure weights within the top 100 institutions is only about one-half as high as that among all academic institutions. These findings lead to two conclusions. First, estimating expenditure weights for the top 100 academic institutions is akin to estimating the same for the population of academic institutions. Two, sample estimates for the top 100 institutions will show much less variance and, therefore, have narrower confidence intervals than sample estimates for the population of academic institutions.

What it true for academic institutions is also true for non-academic institutions. The share of the top 100 non-academic institutions in the total funding received by non-academic institutions is 75 percent (see Table 5.) That is less than the 89 percent controlled by the top 100 academic institutions. Therefore, the weights for the top 100 non-academic institutions differ slightly more from the weights for the population of non-academic institutions. Nonetheless, the differences are slight. Also, the variance of the weights within the top 100 non-academic institutions is dramatically reduced in comparison to the variance within the whole group. Once

[10] An expenditure weight is a ratio of two random variables: expenditures on the category in question, say, wages, and total expenditures. Therefore, variance calculations require the use of the ratio method as described in Yamane (1967, Ch. 13) or Kish (1965, Ch. 6). The variance formula is reproduced in Appendix H. The CV is computed as the ratio of the standard deviation to the mean and is a good way to summarize the spread of observations. In a normal distribution, two-thirds of observations will lie within the boundaries defined by Mean*(1-CV) and Mean*(1+CV). Thus, the higher the CV, the greater the spread in the observed values.

again the suggestion that emerges is that there may be gains to estimating weights for the top 100 institutions as opposed to the population of institutions.[11]

The similarity of large institutions to the population of institutions is revealed from a different perspective in Appendix E. Using the award as the unit of observation and giving all awards the same weight, the analysis in Appendix E shows the following: (1) The higher the number of awards received by an institution, the smaller the deviation of average expenditure shares at that institution from the population mean. Because all mean values are computed as simple averages across awards, the population means, as computed in Appendix E, are not influenced by the number of dollars flowing to large institutions. Thus, the similarity of expenditure shares at large institutions to the population means is not a function of their share in the overall dollars awarded by NIH. (2) The CV of expenditure shares within large institutions closely mimics the CV for the population of awards. Combined with earlier findings, the clear inference is that a survey sample focused on large institutions offers good potential for estimating populations weights in an efficient manner. It appears that institutions receiving a large number of awards from NIH are engaged in a broad spectrum of extramural research that resembles the general characteristics of the population of awards issued by NIH.

Given the variation in expenditure shares across individual observations, the next question is the extent to which the variation is caused by observable characteristics of institutions and awards. Both sample size and sampling error can be reduced if groups of institutions and/or awards can be detected that are internally homogenous with respect to the distribution of expenditure shares. If such groups exist, their defining characteristics can be used to delineate strata for the expenditure survey. In principle, strata might be formed by institution or award characteristics. Examples of institution characteristics that might be used to form strata are size (based on dollars received or number of awards) and kind of organization (institute of higher education, hospital, etc.) Similarly, award characteristics, such as, size of award or type of activity, might be used to form strata. However, since it is feasible to use only the institution as

[11] An alternative to focusing on the top 100 institutions might be to focus on the institutions that collectively account for a given percentage of funding received by all institutions. If the target is 90 percent of funding, 100 academic institutions will suffice but the number of non-academic institutions will increase from 100 to over 300.

the unit of observation, research into the role of award characteristics is primarily just a matter of academic interest.

At an informal level, Appendix C contains tables that show how the expenditure weights vary by major characteristics of institutions and awards. While the data show that expenditure shares do vary with the type of institution or award, those variations have less of an impact than one might expect because the sharpest differences often emerge for characteristics that are associated with only a small share of NIH funding. For example, Table C4 in Appendix C reveals that expenditure shares at Federal institutions are very different from shares at other types of institutions. However, those institutions receive less than one percent of all extramural funding.

A more formal analysis of expenditure weights was also conducted using analysis of variance (ANOVA) and regression analysis to examine the role of institution and award characteristics. Those analyses, conducted for 1991, 1993 and 1995, generated an enormous volume of output that cannot feasibly be presented in its entirety here. Also, since the unit of observation must be the institution and the role of institution size has already been shown to be a critical factor, ANOVA and related analysis are not critical to the design of the sampling strategy. Therefore, the presentation of that analysis is mostly confined to Appendix F.

In brief, ANOVA and regressions analysis did not analyze the variance within all of the expenditure categories that can be estimated with IMPAC data. The analysis instead was focused on the expenditure shares for indirect costs and wages. As shown earlier, these two categories alone account for over two-thirds of total expenditures. ANOVA was also conducted using MDTC as a dependent variable. The reason for that is this variable’s importance to the methodology for computing indirect cost indexes. The institution characteristics used in ANOVA were as follows: region; total dollars awarded; number of awards; kind of organization; and ownership control code. The award characteristics used in ANOVA were as follows: size of award; application type; type of activity; year of support; use of human subjects; and use of animal subjects. Regression analysis was used in part to explore the effect of more detailed characteristics, such as, major component code; scientific class code; and type of department to which the awards is issued.[12]

The major intent of ANOVA was to determine whether mean values of indirect cost shares, wage shares and MDTC shares differ significantly across institutions and awards grouped according to the aforementioned categories. The results showed that virtually all of the institution and award characteristics listed above play a role in shaping the expenditure distribution. An institution effect is to be expected to some extent because indirect cost rates and fringe benefit rates are institution specific. The analysis of institution characteristics found that statistically significant differences in expenditure weights and MDTC exist across institutions grouped by any one of the following characteristics: total dollars awarded; number of awards; kind of organization; ownership control code; and the region in which the institution is located. However, the observed differences are not large in magnitude and arise only when institutions are grouped into fairly broad categories. For example, with respect to wage shares, no significant differences are found between institutions receiving $100m. or more in funding and those receiving $10m. to $24m. But there are significant differences across two broad groups of institutions: those receiving $10m. or more and those receiving $0.5m. to $9m. Similarly, regional differences are not pronounced or consistent from one year to the next.

The analysis of variance also showed that all of the award characteristics listed above are associated with significant differences in expenditure shares and MDTC. For example, projects that use animals have significantly higher indirect cost and supplies shares. In contrast, research with human subjects is associated with lower indirect cost, wage, and supplies shares. Equipment cost shares are much higher for projects in their first year, and wage and indirect cost shares increase with the year of support. With respect to the type of application, supplements and non-competing continuations have significantly higher wage and indirect cost shares. Type of activity matters as well. ANOVA also confirmed that the observed differences in expenditure shares across awards of different sizes are statistically significant.

[12] Definitions of the characteristics of institutions and awards are given in Appendix F.

Regression analysis was used to explore some possibilities that could not be examined with ANOVA alone. As a starting point, expenditure shares (namely, wage shares, indirect cost shares, and modified direct total costs) were regressed on dummy variables representing institution and award characteristics. The main findings were as follows:

1. Institution and award characteristics explain 40 to 60% of the variation in the expenditure shares. The R-squares from the regressions are highest in 1991 and decline somewhat thereafter.
2. Almost all institution and award characteristics are significant.
3. Total dollars received by an institution and the number of awards are often highly significant at the same time.

The regression analysis also confirmed that institution effects played a role in the determination of the expenditure distribution for specific types of awards. Similarly, award characteristics were found to have significant effects within individual institutions.

To summarize, the statistical analysis showed that most institution and award characteristics play a role in shaping the distribution of expenditure shares. This outcome would seem to suggest that a complicated stratification strategy – one that accounts for both institution and award characteristics – might be required to meet the needs of the BRDPI expenditure survey. However, there are several ameliorating factors. First, the test survey has ruled out the use of the award as the unit of observation. Therefore, only institution characteristics matter. Second, it pays to be mindful of the importance of a potential stratifying variable. For example, if one stratifies by kind of organization, it is necessary to take note of the share of NIH funding flowing to each kind of organization. Table D5 in Appendix D shows that institutes of higher education (medical and non-medical) alone account for almost 80 percent of total NIH funding. Thus, building strata by kind of organization simply to make sure a sample of private for-profit businesses is included would be pointless from the point of view of estimating expenditure weights.

The third point to be aware of is the extent to which expenditure shares differ across institutions grouped by potential stratifying variables. Because of the large sample of awards in the IMPAC files, even very small differences in expenditure shares across institution characteristics appear to be statistically significant. But the practical impact on the BRDPI of missing small variations in expenditure shares across institution characteristics is likely to be negligible. It is well known that small variations in expenditure weights matter little in determining the level of a price index. When large differences in expenditure shares do emerge, they often appear for types of institutions that constitute a very small share of total dollars awarded, e.g. Federally owned institutions.

The clearest implication of the analysis of IMPAC data is that institution size, as measured by the share of total dollars awarded, is most important. Categories of awards and institutions that tend to show large differences in expenditure patterns from the overall population generally receive relatively few dollars and are likely to have little influence on estimates of expenditure weights for the BRDPI. It is likely that a focus on institutions receiving a large share of dollars and awards will satisfy the major objectives of the expenditure survey. Focusing on large institutions ensures that the sample represents a good share of total NIH funding. It was shown the top 100 academic and non-academic institutions account for an overwhelming share of funding flowing to each category of institutions. Also, the diversity of awards received by these institutions appears to be a good approximation of the overall characteristics of awards. As shown in the charts in Appendix E, the distribution of expenditure shares within the larger institutions tends to mimic the population distribution. Thus, a sampling strategy focusing on large institutions appears to hold great promise for the BRDPI expenditure survey. The choice of additional stratification variables would appear to depend more on other considerations than statistical concerns. For example, region might be used to stratify the sample if regional estimates of expenditure weights are desired. It must also be kept in mind that the greater the number of stratification variables that are used, the larger the sample size required to populate each strata cell.

5. The Analysis of Institution-Specific Price Indexes

The behavior of price indexes that BEA computes at the level of the institution can affect both the sampling technique and the desired sample size. BEA computes price indexes for wages, fringe benefits and indirect costs separately for each of the 161 academic institutions in its current sample. The institution-specific price indexes are then aggregated using weights derived from IMPAC data, where the weights are the shares of each institution in the total wages, fringe benefits and indirect costs awarded by NIH to the academic institutions in the BEA sample. The indirect cost index for each institution itself is the result of an aggregation process that takes place within individual institutions. In particular, wage and fringe benefits indexes within each institution are aggregated with price indexes for supplies, consultants and travel to yield the modified direct travel cost (MDTC) index. Those components of the MDTC index are combined using weights based on the distribution of expenditures within each institution in the BEA sample. The MDTC index is then used in conjunction with indirect cost rates at individual institutions to compute the indirect cost index for those institutions.[13]

From the point of view of the BRDPI expenditure survey, BEA's current methods raise the following questions: (a) Do the year-to-year price changes differ across institutions grouped by known characteristics? (b) Is the behavior of prices over several years different across institutions? (c) How critical are the weights used by BEA both within and across institutions? To the extent these price indexes are found to show distinct behavior across institutions grouped by size, region, or some other characteristic, those characteristics should be kept in mind when choosing strata for the expenditure survey. On the other hand, if the movement in price indexes is unrelated to institution characteristics, sample designs for the expenditure survey can be developed without regard to price movements. Also, if the price indexes show little or no variation across institutions, a weighted average of institution-specific indexes will be virtually identical to an unweighted average across institutions. It might then be feasible to use alternative methods to weight the institution-specific price indexes for indirect costs, wages and fringe benefits. Of course, that benefit may not amount to much if institution-level detail on the expenditure distribution is still needed for the sake of computing the MDTC index. On the surface it appears that need may only be satisfied by a fairly large sample of institutions. It is necessary, therefore, to determine whether options exist for the estimation of the MDTC index by other methods. If reasonable options do exist and within-institution expenditure detail is no longer required the BRDPI survey could focus on the estimation of the population expenditure distribution and use as small a sample size as is feasible for that purpose.

[13] See the methodology papers issued by BEA and NIH for further details on this aspect of the BRDPI estimation.

The analysis of price movements across institutions was based on a data set provided by BEA that included index levels for wages, fringe benefits, and indirect costs for each of 161 institutions over the period 1979 to 1999.[14] For each institution, the indexes were converted to price relatives measuring year-to-year movements in wages, fringe benefits and indirect costs. An annual average change over the 20-year period was also computed for each institution. ANOVA was then applied to determine if there were any significant differences, within any single year, in the price change (over the previous year) across institutions grouped be selected characteristics. ANOVA was also used to test for significant differences across institutions in the 20-year average annual change in price indexes. The institution characteristics used for the analysis were as follows: major Census region; private/public school; urban/rural location; size based on number of students; size based on number of faculty members; kind of organization; ownership control code; total dollars awarded; and total number of awards. Some of these characteristics, namely, private/public status, urban/rural location, student body and faculty size, were obtained from sources other than the IMPAC file. Those sources are listed in the bibliography.

The analysis of variance showed that, within any single year, the only characteristic that matters is Census region. And even that is not of significance with respect to yearly movements in the indirect cost indexes. With respect to the annual average change over the period 1979-99, it was found that region and ownership control code have some influence on the change in the wage and salary index. In particular, the influence of ownership control code suggests that the accumulated change in wages over time does depend in part on whether or not the institution is state-owned or independent (private.) Fringe benefit movements are independent of any institution characteristics. Accumulated changes in indirect cost indexes were found to depend on the total dollars awarded and number of awards. Since indirect cost rates are negotiated with the Federal government, this finding suggests that negotiating power and outcomes with respect to indirect cost rates depend in part on the extent of the "business" relationship with the government.

[14] The total number of usable observations was 158 after the merger of the BEA data with the IMPAC files to add information on institution characteristics.

Another line of inquiry that was followed was to examine the across-institution variance in year-to-year price movements. If this variance is low, the weights used by BEA to aggregate price relatives across institutions really will not matter much. At first blush, there appeared to be a fairly wide range in price movements across institutions. For example, the wage and salary price relative in 1999 (1998=1) ranged from a low of 0.9989 to a high of 1.1092 across the 158 institutions in the data set. The (simple) average value of the relatives was 1.0437. However, the standard deviation of the observed price relatives was low (0.011), meaning that about two-thirds of the observations lay within a much smaller range (1.0327 to 1.0547.)[15] Also, within any single year, it was found that the standard error of the simple average of price relatives was quite low and the resulting 95% confidence interval was fairly narrow. In other words, for all three types of price indexes, it was possible to be fairly certain that the true average of the price relatives lay within a very tight range of the computed simple average of the price relatives.[16] This suggests that weights are not critical to the BEA methodology for computing wage and salary, fringe benefits, and indirect cost indexes.

The possibility that institution-level weights are not critical to the BEA methodology for computing wage, fringe benefit and indirect cost indexes was tested by re-estimating those indexes from the BEA data using a variety of weighting strategies. The weighting strategies tried with the BEA price index data were as follows: (1) Use alternative weights based on the distribution of total dollars awarded or total direct costs across institutions. (2) Group institutions by major characteristics and use weights only at the group level. With regard to this strategy, the following groups were tried: four major regional groups (north-east, south, mid-west, west), kind of organization (medical or non-medical), ownership type (state, local, independent), total dollars

---

[15] In a normal distribution, two-thirds of observations lie within an interval defined by adding and subtracting the standard deviation from the mean.

[16] Another approach that was tried was to compute similar data on standard deviations and standard errors within individual institutions across time. The standard errors and confidence intervals for price relatives within institutions over time were found to be much larger than across institutions within a single point of time. However, from the point of view of the BRDPI, the question is whether institutions showed varying tendencies over time. The answer to that is there do seem to be differences across institutions with respect to the price movements they show over time. However, as revealed by the ANOVA analysis, these differences do not seem to correlate with any of the known characteristics of individual institutions.

awarded (seven categories) and total number of awards (six categories). (3) Use no weights, i.e. compute simple averages of the price indexes across institutions.

Partial results from the use of the first two methods are shown in Tables 8, 9 and 10 below. More detailed results are presented in Appendix G. Table 8 shows the levels of the salary and wages price indexes computed using four different weighting strategies. The BEA Salary and Wages Index in Table 8 is the one computed by BEA using institution-level weights derived from the distribution of wage expenditures across institutions. The first two alternative indexes use institution-level weights, but instead of relying on the distribution of wage expenditures they use the across-institution distribution of total dollars awarded and total direct costs respectively. The reason for testing these alternatives is that institution-level data on total dollars awarded and total direct costs are still available from the IMPAC files. The next two alternative indexes use "group-level" weights. This means that institutions were grouped into categories based on total dollars awarded or the number of awards received. Within these categories, institution-specific indexes were (geometrically) averaged without the use of weights. However, weights based on the distribution of wage expenditures were then applied to aggregate across groups. The reason for testing this method is that, depending upon sample size, the BRDPI expenditure survey may be able to yield sufficient information to estimate wage-expenditure-based weights for groups of institutions.

The results in Table 8 show that, for the salary and wages index, the choice of weighting strategy appears to be virtually irrelevant. Indexes computed using any of the alternative strategies are identical to those computed by BEA. The BEA salary and wages index increased 192.3% between 1979 and 1999 at an average annual rate of 5.51%. The trends in the alternative indexes are no different. The limited variance in the salary and wages price index across institutions is one reason for this result. Another reason is that weights based on the across-institution distribution of wages, total direct costs and total dollars awarded are nearly perfectly correlated.[17]

[17] The correlation coefficients are between 0.99 and 1.

Table 8
Salary and Wages Index Estimated by BEA for Academic Institutions Compared to Indexes Derived with Alternative Weighting Strategies

Year	*BEA Salary & Wages Index*	*Alternative Index with Institution-Level Weights* Total Dollars Awarded (1)	Total Direct Costs (2)	*Alternative Index with Group-Level Weights* Total Dollars Awarded (3)	Number of Awards (4)	BEA Index less (1)	BEA Index less (2)	BEA Index less (3)	BEA Index less (4)
1979	42.17	42.17	42.20	42.18	42.25	-0.01	-0.04	-0.02	-0.08
1980	45.66	45.66	45.70	45.67	45.68	0.00	-0.04	-0.01	-0.02
1981	50.00	50.00	50.05	50.01	50.01	0.00	-0.05	0.00	-0.01
1982	54.76	54.76	54.81	54.75	54.76	0.00	-0.05	0.01	0.00
1983	57.95	57.96	57.99	57.97	58.01	0.00	-0.04	-0.02	-0.05
1984	61.41	61.40	61.43	61.40	61.50	0.00	-0.03	0.01	-0.10
1985	65.59	65.58	65.61	65.59	65.65	0.01	-0.02	0.00	-0.05
1986	69.32	69.31	69.36	69.33	69.30	0.01	-0.04	-0.01	0.02
1987	73.57	73.57	73.61	73.58	73.45	0.00	-0.04	-0.01	0.12
1988	77.78	77.78	77.82	77.80	77.70	0.00	-0.04	-0.02	0.08
1989	81.64	81.63	81.67	81.67	81.64	0.01	-0.03	-0.03	0.00
1990	86.49	86.47	86.52	86.53	86.57	0.02	-0.03	-0.04	-0.08
1991	91.20	91.18	91.23	91.22	91.29	0.02	-0.03	-0.02	-0.09
1992	96.08	96.10	96.12	96.07	96.00	-0.02	-0.04	0.01	0.08
1993	100.00	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00
1994	102.80	102.80	102.80	102.79	102.89	-0.01	0.00	0.01	-0.09
1995	106.25	106.27	106.24	106.27	106.31	-0.02	0.01	-0.02	-0.06
1996	108.95	108.96	108.94	109.00	109.15	-0.01	0.00	-0.05	-0.21
1997	112.97	113.01	112.96	113.02	113.10	-0.03	0.02	-0.05	-0.12
1998	118.02	118.06	118.01	118.05	117.97	-0.03	0.01	-0.02	0.06
1999	123.26	123.27	123.24	123.30	123.17	-0.01	0.02	-0.04	0.09
Percentage Change:									
1979 to 1999	192.3%	192.3%	192.0%	192.3%	191.5%	0.01%	0.30%	0.01%	0.79%
Avg. Annual	5.51%	5.51%	5.50%	5.51%	5.50%	0.00%	0.01%	0.00%	0.01%

Note: See text and Appendix G for further explanation and methodological details.
Source: Joel Popkin and Company using IMPAC data and BEA data on the BRDPI.

Table 9
Fringe Benefits Indexes Estimated by BEA for Academic Institutions Compared to Indexes Derived with Alternative Weighting Strategies

		Alternative Index with Institution-Level Weights		*Alternative Index with Group-Level Weights*					
Year	*BEA Fringe Benefits Index*	Total Dollars Awarded (1)	Total Direct Costs (2)	Total Dollars Awarded (3)	Number of Awards (4)	BEA Index less (1)	BEA Index less (2)	BEA Index less (3)	BEA Index less (4)
1979	35.98	36.52	36.73	36.27	35.31	-0.54	-0.75	-0.28	0.68
1980	39.57	40.21	40.45	39.93	38.79	-0.64	-0.88	-0.36	0.77
1981	44.80	45.66	45.98	45.34	44.00	-0.86	-1.18	-0.55	0.80
1982	50.57	51.42	51.73	51.14	50.04	-0.86	-1.16	-0.57	0.53
1983	54.44	55.20	55.52	54.96	53.80	-0.76	-1.08	-0.52	0.64
1984	59.75	60.50	60.84	60.23	58.87	-0.75	-1.08	-0.48	0.89
1985	64.51	65.05	65.30	64.93	64.15	-0.54	-0.80	-0.42	0.36
1986	67.18	67.82	68.04	67.91	67.05	-0.64	-0.87	-0.73	0.12
1987	70.50	71.18	71.45	71.22	70.69	-0.69	-0.96	-0.72	-0.19
1988	76.58	77.28	77.57	77.28	76.61	-0.70	-0.99	-0.70	-0.03
1989	81.21	82.06	82.38	81.82	81.00	-0.86	-1.17	-0.61	0.21
1990	86.48	87.22	87.48	86.90	86.28	-0.74	-1.00	-0.43	0.20
1991	92.21	92.90	93.16	92.58	92.27	-0.69	-0.95	-0.37	-0.06
1992	97.74	98.29	98.53	98.07	97.87	-0.54	-0.78	-0.33	-0.12
1993	100.00	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00
1994	110.27	110.81	111.00	110.62	109.56	-0.54	-0.73	-0.36	0.71
1995	115.35	115.90	116.10	115.78	114.52	-0.54	-0.75	-0.43	0.83
1996	121.09	122.29	122.69	122.02	120.14	-1.20	-1.61	-0.94	0.95
1997	124.31	125.45	125.80	125.20	123.29	-1.15	-1.49	-0.89	1.02
1998	127.02	128.48	128.97	128.00	126.00	-1.46	-1.95	-0.98	1.02
1999	131.79	133.30	133.80	132.93	130.66	-1.51	-2.01	-1.14	1.13
Percentage Change:									
1979 to 1999	266.3%	265.0%	264.3%	266.5%	270.1%	1.28%	1.97%	-0.28%	-3.82%
Avg. Annual	6.71%	6.69%	6.68%	6.71%	6.76%	0.02%	0.03%	0.00%	-0.06%

Note: See text and Appendix G for further explanation and methodological details.

Source: Joel Popkin and Company using IMPAC data and BEA data on the BRDPI.

Table 10
Indirect Cost Indexes Estimated by BEA for Academic Institutions Compared to Indexes Derived with Alternative Weighting Strategies

Year	*BEA Indirect Cost Index*	*Alternative Index with Institution-Level Weights*	*Alternative Index with Group-Level Weights*				
		Total Dollars Awarded	Total Dollars Awarded	Number of Awards	BEA Index less (1)	BEA Index less (2)	BEA Index less (3)
		(1)	(2)	(3)			
1979	34.61	34.75	34.71	35.09	-0.14	-0.10	-0.48
1980	39.17	39.21	39.24	39.55	-0.04	-0.07	-0.38
1981	44.05	43.67	43.69	43.92	0.38	0.36	0.13
1982	48.89	48.76	48.99	48.90	0.13	-0.10	-0.01
1983	53.86	53.59	53.95	53.68	0.27	-0.08	0.18
1984	60.08	59.77	59.92	59.54	0.31	0.17	0.55
1985	63.75	63.45	63.70	63.35	0.31	0.05	0.40
1986	67.45	67.16	67.54	66.99	0.28	-0.09	0.46
1987	72.12	72.01	72.24	71.87	0.11	-0.12	0.25
1988	75.72	75.93	76.17	75.78	-0.21	-0.45	-0.06
1989	80.42	80.71	80.99	80.49	-0.29	-0.56	-0.07
1990	86.69	87.05	87.23	87.07	-0.36	-0.54	-0.38
1991	91.58	91.81	91.98	92.07	-0.23	-0.40	-0.49
1992	96.79	97.00	97.03	97.04	-0.21	-0.25	-0.25
1993	100.00	100.00	100.00	100.00	0.00	0.00	0.00
1994	104.33	104.51	104.57	104.50	-0.18	-0.24	-0.17
1995	107.90	108.49	108.38	108.32	-0.59	-0.48	-0.42
1996	110.33	111.13	111.04	110.96	-0.80	-0.70	-0.63
1997	113.01	113.70	113.69	113.63	-0.70	-0.68	-0.62
1998	117.74	118.60	118.61	118.26	-0.86	-0.87	-0.53
1999	121.58	122.74	122.87	122.53	-1.16	-1.29	-0.95
Percentage Change:							
1979 to 1999	251.3%	253.2%	254.0%	249.2%	-1.96%	-2.72%	2.07%
Avg. Annual	6.48%	6.51%	6.52%	6.45%	-0.03%	-0.04%	0.03%

Note: See text and Appendix G for further explanation and methodological details.
Source: Joel Popkin and Company using IMPAC data and BEA data on the BRDPI.

Tables 9 and 10 present the results for the fringe benefits and the indirect cost indexes. Strictly speaking, there is no need to test alternatives for the indirect cost index because data on indirect costs for individual institutions are still recorded in the IMPAC files. However, alternative methods were tested for the sake of completeness. One exception is that the method of deriving weights based on the distribution of direct costs is not applied in the case of the indirect cost index. The results for the fringe benefits and indirect cost indexes are also very clear, namely, a number of alternatives to the current BEA method are available to aggregate institution-specific indexes across the 150 plus institutions in the BEA sample.[18] Of the various methods reported in Tables 8 to 10 and in Appendix G the most appealing is the method that uses institution-level weights based on the distribution of total dollars awarded. This information remains readily available from the IMPAC files. Compared to the BEA indexes computed over the 1979 to 1999 time period, this alternative would have had no impact on the average annual rate of change in the salary and wages index. The fringe benefits index would have increased faster by 0.02 percentage points per year and the indirect cost index would have increased slower by 0.03 percentage points per year.

The last issue with respect to institution-specific price indexes is the estimation of the MDTC index for each academic institution. As explained at the beginning of this section, that process currently depends upon knowing the expenditure distribution within each of the 161 institutions in the BEA sample. The only way to preserve a semblance of the current methodology under the BRDPI expenditure survey is to have a relatively large sample of 50 to 75 academic institutions. BEA could then either scale back its own sample from 161 to match the sample from the expenditure survey or the survey data could be used to impute the expenditure distribution for institutions in the BEA sample but not in the survey sample. If the expenditure survey sample is small, BEA would need to apply alternative methods for the estimation of the MDTC index.

[18] The distributions of total dollars, direct costs, indirect costs and fringe benefits across institutions are also very highly correlated. The correlation coefficients are 0.97 and higher. As was the case with the salary and wages index, the high correlation in the weights combined with the low variances of the indirect cost and fringe benefits indexes across institutions is responsible for the very similar results obtained from the various methods.

One alternative for BEA is to substitute estimates of the population expenditure shares for institution-specific expenditure shares within each institution. Simulations with this procedure showed that between 1979 and 1999 this method would have led to a smaller increase in the MDTC index for 73 institutions and a larger increase in the remaining 85 institutions.[19] In 39 out of the 158 institutions included in the simulations, the average annual increase in the MDTC index derived using the alternative method differed from the BEA method by 0.1 percentage point or more in absolute value. The largest absolute difference in the average annual increase was 0.31 percentage points. With respect to the total change between 1979 and 1999, the difference between the BEA method and the alternative method exceeded 5 percentage points in absolute value for 48 institutions. In only 9 cases did the twenty-year difference exceed 10 percentage points in absolute value, with the largest difference being –19.34 percentage points.

While the substitution of population shares for institution-level shares could lead to significant differences in the MDTC index for some institutions, the differences are very small for the large majority of institutions. It is also important to keep in mind that it is the aggregate difference across all institutions that matters. Smaller increases in the MDTC index in some institutions are offset by higher increases in other institutions. One way to look at this issue is to take a simple average of the differences across institutions. The results show that the mean difference in the average annual rate of change is only 0.01 percentage point and the mean difference in the 20-year accumulated change is only 0.58 percentage points.[20] Thus, the overall bias in the MDTC index and, by inference, the indirect cost index is likely to be small. Since indirect costs have a weight of about 0.3 in the BRDPI for academic institutions, which translates to a weight of only 0.2 in the overall BRDPI, the bias in the BRDPI itself will be even smaller and of little, if any, significance.

[19] See Appendix G for detailed results.
[20] The absolute difference in the average annual rate of change, when averaged across the 158 institutions, is 0.07 percentage points. Similarly, the absolute difference in the twenty-year change averages to 3.96 percentage points across all institutions.

6. Tests of Sampling Strategies

The availability of IMPAC data for 1991, 1993 and 1995 makes it possible to draw samples using several different strategies and compare the resulting estimates of expenditure weights with the known population weights for those years. The results of these tests can then form the basis of developing a sample for the final version of the BRDPI expenditure survey. The unit of observation for these tests is the institution because that will also be the case with the BRPDI expenditure survey. Although the test samples were drawn for 1991, 1993 and 1995, this section reports results only for 1993. Results for other years are very similar and are presented in Appendix H.

The experimental samples were designed to estimate weights for the population of institutions and for the smaller population of the top 100 institutions. As shown in Section 4 above, the top 100 academic institutions account for 90 percent of the total funding received by academic institutions and the top 100 non-academic institutions receive 75 percent of the total dollars awarded to that category of institutions. Consequently, the expenditure distribution for the top 100 is virtually identical to that for the overall population and the former group can be thought of an alternative universe from which samples could be drawn to satisfy the needs of the BRDPI. An advantage of focusing on the top 100 is a sharply reduced variance in the expenditure weights across institutions.

The focus of the test samples was on uncovering the range of point estimates that might emerge from the variety of sampling strategies that could potentially be used in the final survey. Given the various estimates of the expenditure weights, there is a need to choose one or two strategies as lead candidates for the BRDPI expenditure survey. The choice depends upon both statistical criteria and practical considerations. The principal statistical criterion is a low variance of the sample estimate. Since variance can always be reduced be picking a larger sample size there is a need to establish a criterion that helps determine the level of tolerance for the variance. One method of making that determination is to use the sample weights to re-estimate the extramural component of the BRDPI. If the BRDPI is not too sensitive to sample design, one can afford to have higher tolerance for variance in the sample estimates. Practical considerations that must be kept in mind include the potential need for OMB clearance, respondent burden, and the cost of the survey. All of these factors argue in favor of keeping the sample size as small as possible.

The principal finding of the sampling experiments is that accurate data can be obtained from a variety of strategies using a relatively small number of institutions. The samples could be either probability samples (e.g. stratified random samples) or non-probability samples (e.g. the nine largest institutions based on total dollars awarded.) In short, there are a number of strategies that yield accurate estimates of expenditure weights with low variance and BRDPI levels that closely track BEA estimates. One can thank the skewed nature of the distribution of NIH funding for this result. As noted earlier in the report, a relatively small number of large institutions account for the majority of extramural funding as well as a large (and representative) portion of awards. Since the implicit goal of the expenditure survey is to cover dollars, not institutions, a small sample of institutions can yield the necessary coverage in light of the skewed distribution of funding. In particular, a sample of 9 to 18 large academic institutions and 9 to 27 large non-academic institutions will be sufficient. The small sample sizes will mean low respondent burden and modest survey costs. OMB clearance can also be avoided by spreading the survey over two to three years so that no more than nine institutions are contacted in any single year.

6.1. Estimates of Expenditure Weights from Experimental Samples

This section presents expenditure weights estimated from a range of samples drawn from the IMPAC data. The variance estimates shown in the next section address the issue of what might result from the repeated drawing of the same types of samples. Section 6.3 rounds off the analysis with a presentation of extramural BRDPI estimates derived using sample weights as opposed to the population weights used by BEA.

In principle, the IMPAC data could be used to test a large number of sampling strategies utilizing as many stratification variables as possible. Data on the following major institution characteristics are contained in the IMPAC files: region; total dollars awarded; number of awards; kind of organization; and ownership control code. The required division of the sample into academic and non-academic institutions indirectly accounts for kind of organization. In deciding which of the remaining characteristics should be used to form strata, it is important to note the number of academic and non-academic institutions receiving extramural funding from the NIH. At the present time, there are approximately 475 academic and 1,250 non-academic institutions receiving funding from the NIH. Given the relatively small number of academic institutions, a complicated stratification plan for this group could quickly begin to exhaust the total number of institutions. For example, four Census regions, five size classes based on total dollars received, and a division by public or private ownership leads to 40 cells that must be populated with the available total of 475 academic institutions. In addition, as overall sample size grows with the number of strata, so does respondent burden.

It is desirable, therefore, to use only a limited number of stratification variables. Other than the required split by academic/non-academic status, the most obvious stratification variable is total dollars awarded to an institution.[21] Given the concentration of NIH funding within both academic and non-academic institutions, a focus on institution size guarantees that a relatively small sample can cover a large proportion of the total dollars obligated by NIH. The institutions that dominate in terms of funding also receive a large and representative share of awards.

The major sampling strategies that were tested and reported on below are as follows:

Random Sampling: Simple random samples drawn at the rate of one in five, one in ten and one in twenty institutions. See Appendix H for details on the sampling technique.

Stratified Random Sampling – Academic Institutions: Two types of strata were formed for academic institutions. In the *Two Strata* case, academic institutions were divided into two groups – those receiving \$5m. or more in funding and those receiving less than \$5m. In the *Five Strata* case, academic institutions were divided into the following five groups based on total dollars awarded: \$5m. or more; \$1-4m.; \$0.5-0.9m.; \$0.25-0.4m.; less than \$0.25m. The sampling rate was highest in the strata with the largest institutions. The precise sampling rate within each stratum varied depending on the objective for the total sample size. Details are discussed in Appendix H.

Stratified Random Sampling – Non-academic Institutions: In addition to the *Two Strata* and *Five Strata* strategies defined above, non-academic institutions were also split

[21] An alternative to stratifying by size is to design a random sample where the probability of selection increases with institution size.

into *Ten Strata* based on total dollars awarded.[22] The reason for forming ten strata is that non-academic institutions are a more heterogeneous group than academic institutions. The ten strata are as follows: institutions receiving $25m. or more in funding; $5-24m.; $2.5-4m.; $1-2.4m.; $0.75-0.9m.; $0.5-0.74m.; $0.35-0.49m.; $0.25-0.34m.; $0.1-0.24m.; less than $0.1m. Once again, the sampling rate within each stratum varied depending on the desired sample size and the details are discussed in Appendix H.[23]

By Region: The top two institutions from each of the four major Census regions plus one other large institution were picked on the basis of total dollars awarded.

By Number of Awards: The top nine and top twenty institutions ranked by the number of awards received.

By Total Dollars Awarded: The top nine and top twenty institutions ranked by total dollars awarded.

Random Samples from the Top 100 Institutions: Simple random samples of 9, 18 and 27 each from the Top 100 academic institutions and the Top 100 non-academic institutions.

The results from the sampling experiments are shown in Tables 11 to 14. Expenditure weights for the population of institutions are also shown in all tables. As explained in Section 4.2, expenditure weights are computed using the plutocratic method, i.e. estimates of the population expenditure shares are weighted averages of institution-level shares where the weights are based on the distribution of total dollars awarded across institutions. For academic institutions, the weights for the direct costs components are renormalized so that their sum is equal to direct costs, less other costs. The other cost category is excluded because it is not a part of the extramural BRDPI. The indirect cost category does not appear for non-academic institutions for the same reason. The expenditure categories of principal interest are salary and wages, fringe benefits and supplies. Those three categories account for over 90 percent of total direct costs, less other costs. Sample estimates of indirect costs are not critical because

[22] A four strata strategy was also tried but is not reported on for the sake of brevity.

[23] In addition to size, one may wish to form strata based on the public or private ownership of institutions. That is because indirect cost rates are known to differ systematically across these two types of institutions. Evidence to that effect is presented in Appendix C, Table C4, where the relevant comparison is between state-owned and independent institutions. However, the observed differences in weights are relatively small. Further, initial testing showed that stratifying by academic/non-academic status, total dollars awarded and public/private status would lead to a number of empty or near empty cells for non-academic institutions. That is because non-academic institutions are overwhelmingly private and those institutions also account for 95 percent of the total dollars awarded to non-academic institutions (see Appendix D, Table D4.) With respect to academic institutions, they are nearly equally split between public and private ownership along most dimensions. As a result, any sampling strategy leads to an adequate coverage of both public and private institutions. For example, the top 100 academic institutions include 59 public and 41 private institutions. Any random sample from this group is likely to have a good representation of both types of institutions.

population weights for indirect costs can still be computed from the IMPAC files.[24] Comparing the point estimates from the samples with the population weights reveals that all of the tested sampling strategies are promising.

Table 11 reports on the random samples that were each drawn three times from the 1993 IMPAC files.[25] The results for both academic and non-academic institutions appear quite satisfactory. The point estimates for expenditure weights lie within reasonable range of the population means, at least for the three sample drawings underlying the data in Table 11. One reason the random samples perform well is that the 10 and 20 percent sampling rates are quite high. The 5 percent random samples also lead to a fairly large sample of institutions, about 75 to 100 academic and non-academic institutions combined. Based on the results of the test survey, that implies a total respondent burden of up to 900 hours. Also, OMB approval would be required before conducting a survey of this size. The 75 to 100 institutions that comprise a 5 percent sample cover just about 5 percent of total dollars awarded. As shown in Table 5, the top 9 institutions alone command over a 20 percent share of total dollars awarded. Thus, it would be more efficient to focus on large institutions using strategies other than simple random sampling.

The results obtained from the stratified samples of various sizes are presented in Table 12. Estimates from those samples also lie within small ranges around the population means. In comparison to the simple random samples, the results from the stratified samples appear to lie closer to the population weights. Sample size also does not seem to make much difference. These points are confirmed by the more formal variance analysis presented in the next section.

---

[24] It should, however, be kept in mind that weights estimated from IMPAC data derive from planned expenditures whereas the expenditure survey reports on actual expenditures. Therefore, the weight for the indirect cost category estimated from the IMPAC data could be different from the weight that might be derived from a survey.

[25] Sample sizes differ within a given sampling rate (e.g. one-in-five) due to the particular technique used to draw the random sample. See Appendix H for details. The sampling was also done in a way so that, within each strategy, there was no overlap across the three different samples. The 5, 10 and 20 percent random samples of institutions also lead to roughly the same coverage of total dollars awarded.

Table 11
Expenditure Shares Derived from Random Samples of Institutions from the IMPAC Files, 1993

		One in Five Random Samples			**One in Ten Random Samples**			**One in Twenty Random Samples**		
Academic Institutions	**Population**	*Sample 1*	*Sample 2*	*Sample 3*	*Sample 1*	*Sample 2*	*Sample 3*	*Sample 1*	*Sample 2*	*Sample 3*
No. of Observations	472	88	99	109	42	58	50	17	22	21
Indirect Cost Share	0.2967	0.3004	0.2970	0.2856	0.2966	0.2995	0.2957	0.2978	0.2983	0.3165
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6362	0.6433	0.6327	0.6253	0.6631	0.6250	0.6365	0.6680	0.6504	0.6290
Fringe benefits	0.1554	0.1456	0.1615	0.1627	0.1269	0.1657	0.1663	0.1269	0.1596	0.1540
Consultants	0.0061	0.0063	0.0062	0.0056	0.0054	0.0058	0.0063	0.0048	0.0060	0.0048
Equipment	0.0284	0.0316	0.0287	0.0302	0.0360	0.0304	0.0281	0.0322	0.0267	0.0303
Supplies	0.1401	0.1399	0.1369	0.1452	0.1405	0.1434	0.1313	0.1391	0.1320	0.1498
Travel	0.0184	0.0177	0.0179	0.0186	0.0176	0.0179	0.0189	0.0162	0.0198	0.0191
Patient care	0.0137	0.0134	0.0152	0.0113	0.0102	0.0110	0.0113	0.0124	0.0047	0.0127
Alterations	0.0017	0.0022	0.0008	0.0010	0.0003	0.0008	0.0014	0.0003	0.0009	0.0003

		One in Five Random Samples			**One in Ten Random Samples**			**One in Twenty Random Samples**		
Non-academic Institutions	**Population**	*Sample 1*	*Sample 2*	*Sample 3*	*Sample 1*	*Sample 2*	*Sample 3*	*Sample 1*	*Sample 2*	*Sample 3*
No. of Observations	1150	256	245	225	125	114	138	74	66	47
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6423	0.6253	0.6502	0.6538	0.6136	0.6571	0.6709	0.6633	0.6702	0.6580
Fringe benefits	0.1483	0.1486	0.1451	0.1550	0.1530	0.1495	0.1370	0.1518	0.1329	0.1639
Consultants	0.0176	0.0167	0.0219	0.0123	0.0168	0.0127	0.0197	0.0180	0.0226	0.0245
Equipment	0.0291	0.0332	0.0270	0.0241	0.0365	0.0248	0.0263	0.0254	0.0228	0.0284
Supplies	0.1250	0.1447	0.1150	0.1165	0.1558	0.1110	0.1133	0.1160	0.1120	0.1021
Travel	0.0232	0.0202	0.0228	0.0236	0.0201	0.0278	0.0195	0.0224	0.0198	0.0210
Patient care	0.0139	0.0114	0.0180	0.0145	0.0042	0.0168	0.0132	0.0032	0.0196	0.0021
Alterations	0.0005	0.0000	0.0000	0.0002	0.0000	0.0002	0.0000	0.0000	0.0002	0.0000

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 12
Expenditure Shares Derived from Stratified Samples of Institutions from the IMPAC Files, 1993

		Stratified Samples									
Academic Institutions	**Population**	**2-STRATA**					**5-STRATA**				
No. of Observations	472	25	51	66	107	208	28	50	82	110	216
Indirect Cost Share	0.2967	0.3213	0.2887	0.2823	0.2944	0.2964	0.3179	0.2891	0.2820	0.2950	0.2970
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6362	0.6308	0.6387	0.6394	0.6450	0.6360	0.6268	0.6406	0.6373	0.6445	0.6355
Fringe benefits	0.1554	0.1630	0.1497	0.1544	0.1512	0.1554	0.1642	0.1496	0.1550	0.1510	0.1552
Consultants	0.0061	0.0053	0.0068	0.0065	0.0060	0.0061	0.0064	0.0060	0.0063	0.0058	0.0061
Equipment	0.0284	0.0244	0.0340	0.0299	0.0276	0.0288	0.0258	0.0318	0.0312	0.0278	0.0289
Supplies	0.1401	0.1428	0.1412	0.1343	0.1375	0.1397	0.1424	0.1427	0.1346	0.1385	0.1404
Travel	0.0184	0.0163	0.0192	0.0204	0.0188	0.0186	0.0170	0.0187	0.0205	0.0186	0.0185
Patient care	0.0137	0.0168	0.0088	0.0129	0.0131	0.0138	0.0167	0.0088	0.0129	0.0132	0.0137
Alterations	0.0017	0.0006	0.0015	0.0022	0.0008	0.0016	0.0006	0.0018	0.0022	0.0008	0.0016

		Stratified Samples											
Non-academic Institutions	**Population**	**2-STRATA**					**5-STRATA**					**10-STRATA**	
No. of Observations	1150	26	53	75	144	277	27	52	109	116	206	52	145
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6423	0.6522	0.6279	0.6513	0.6359	0.6355	0.6809	0.6332	0.6509	0.6434	0.6467	0.6330	0.6371
Fringe benefits	0.1483	0.1504	0.1431	0.1546	0.1524	0.1495	0.1312	0.1542	0.1499	0.1459	0.1487	0.1375	0.1518
Consultants	0.0176	0.0078	0.0247	0.0212	0.0130	0.0195	0.0095	0.0170	0.0172	0.0171	0.0169	0.0186	0.0215
Equipment	0.0291	0.0318	0.0514	0.0267	0.0323	0.0322	0.0207	0.0322	0.0289	0.0303	0.0273	0.0306	0.0267
Supplies	0.1250	0.1229	0.1158	0.1111	0.1314	0.1277	0.1180	0.1276	0.1256	0.1285	0.1243	0.1309	0.1187
Travel	0.0232	0.0156	0.0229	0.0269	0.0210	0.0223	0.0165	0.0173	0.0190	0.0220	0.0217	0.0301	0.0277
Patient care	0.0139	0.0183	0.0132	0.0081	0.0142	0.0129	0.0222	0.0176	0.0083	0.0126	0.0138	0.0192	0.0155
Alterations	0.0005	0.0011	0.0010	0.0000	0.0000	0.0005	0.0011	0.0010	0.0001	0.0002	0.0006	0.0000	0.0010

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 13
Expenditure Shares Derived from Non-Probability Samples of Large Institutions from the IMPAC Files, 1993

			By Number of Awards		By Total Dollars Awarded	
Academic Institutions	**Population**	**By Region**	*Top 9*	*Top 20*	*Top 9*	*Top 20*
No. of Observations	472	9	9	20	9	20
Indirect Cost Share	0.2967	0.3048	0.2966	0.2978	0.3030	0.2995
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6362	0.6480	0.6415	0.6397	0.6409	0.6394
Fringe benefits	0.1554	0.1548	0.1621	0.1590	0.1644	0.1590
Consultants	0.0061	0.0054	0.0056	0.0054	0.0058	0.0056
Equipment	0.0284	0.0236	0.0219	0.0245	0.0209	0.0248
Supplies	0.1401	0.1349	0.1318	0.1364	0.1305	0.1354
Travel	0.0184	0.0170	0.0191	0.0181	0.0188	0.0182
Patient care	0.0137	0.0148	0.0157	0.0153	0.0159	0.0150
Alterations	0.0017	0.0013	0.0024	0.0017	0.0027	0.0027

			By Number of Awards		By Total Dollars Awarded	
Non-academic Institutions	**Population**	**By Region**	*Top 9*	*Top 20*	*Top 9*	*Top 20*
No. of Observations	1150	9	9	20	9	20
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6423	0.6558	0.6542	0.6515	0.6542	0.6425
Fringe benefits	0.1483	0.1452	0.1526	0.1506	0.1526	0.1530
Consultants	0.0176	0.0046	0.0044	0.0058	0.0044	0.0056
Equipment	0.0291	0.0222	0.0205	0.0212	0.0205	0.0219
Supplies	0.1250	0.1420	0.1415	0.1387	0.1415	0.1468
Travel	0.0232	0.0150	0.0126	0.0142	0.0126	0.0134
Patient care	0.0139	0.0152	0.0141	0.0169	0.0141	0.0157
Alterations	0.0005	0.0000	0.0000	0.0011	0.0000	0.0011

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 14
Expenditure Shares Derived from Top 100 Samples of Institutions from the IMPAC Files, 1993

			Random Samples		
Academic Institutions	**Population**	**Top 100**	*Sample 1*	*Sample 2*	*Sample 3*
No. of Observations	472	100	9	18	27
Indirect Cost Share	0.2967	0.2989	0.2965	0.2954	0.2899
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6362	0.6360	0.6430	0.6252	0.6379
Fringe benefits	0.1554	0.1569	0.1580	0.1535	0.1533
Consultants	0.0061	0.0056	0.0051	0.0049	0.0056
Equipment	0.0284	0.0265	0.0263	0.0321	0.0274
Supplies	0.1401	0.1404	0.1350	0.1523	0.1448
Travel	0.0184	0.0179	0.0182	0.0178	0.0172
Patient care	0.0137	0.0149	0.0143	0.0137	0.0122
Alterations	0.0017	0.0018	0.0000	0.0005	0.0017

			Random Samples		
Non-academic Institutions	**Population**	**Top 100**	*Sample 1*	*Sample 2*	*Sample 3*
No. of Observations	1150	100	9	18	27
Direct Cost Shares:	1.0000	1.0000	1.0000	1.0000	1.0000
Salary and wages	0.6423	0.6428	0.6421	0.6332	0.6305
Fringe benefits	0.1483	0.1534	0.1721	0.1721	0.1590
Consultants	0.0176	0.0106	0.0090	0.0142	0.0063
Equipment	0.0291	0.0222	0.0192	0.0160	0.0248
Supplies	0.1250	0.1361	0.1311	0.1109	0.1507
Travel	0.0232	0.0194	0.0252	0.0330	0.0171
Patient care	0.0139	0.0148	0.0011	0.0204	0.0114
Alterations	0.0005	0.0007	0.0001	0.0000	0.0000

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 13 displays the findings obtained from the non-probability samples. Despite the small size of these samples, the results are excellent and confirm that large institutions mimic the characteristics of the population expenditure distribution. Based on the point estimates in Table 13, there also appears to be little gain from increasing the sample size from the nine largest to the twenty largest institutions. And, as revealed by the regional sample, equally good results are obtained without necessarily having to select only the top nine institutions. Sampling from the Top 100 institutions is an alternative way of focusing attention on large institutions. The results are shown in Table 14. Because the 100 largest institutions make up a fairly homogenous group, random samples as small as nine institutions are seemingly sufficient to provide good results.

6.2. Estimates of Sampling Variance

The estimates of expenditure weights reported in the preceding section are based on only one to three sample drawings using any given strategy. Due to the underlying variance in the data, repeated applications of any strategy will yield different estimates of the expenditure shares. Therefore, an important criterion for choosing among the sampling strategies is the variance of the expenditure shares. The smaller the variance, the greater the reliability of a sampling strategy. With the availability of the IMPAC files, the variance estimates can be derived from the underlying population data for any hypothetical sample size. Of course, variance estimates can only be derived for probability samples. In other words, the variance criterion cannot be used to decide on the efficacy of the non-probability samples, such as, the top 9 or the top 20 institutions. With regard to the top 100 institutions, one can redefine the population to consist only of the top 100 and compute the variance of the samples in Table 14 using the standard formulas. However, care should be taken not to compare the variance estimates for sample estimates from the top 100 institutions with sample estimates from the entire population.

Variance estimates for simple random samples and stratified random samples were derived using formulas applicable to ratios of two random variables. Those formulas are given in Yamane (1967) and are also reproduced in Appendix H.[26] Variance estimates were derived for

[26] See Chapter 13 in Yamane. Also, see Kish (1965), Chapter 6.

simple random and stratified random samples of size 10, 25, 50, 100, 150 and 200 from the IMPAC population data. This is a more general approach than confining the variance calculations to sample sizes equal to those reported in Tables 11 to 13. Separate procedures were followed for the sample of the Top 100 institutions. Assuming the Top 100 to represent the population at large, variance estimates were computed for random samples of 9, 18 and 27 drawn from the "population" of 100.

The variance estimates were then used to compute confidence intervals for the expenditure weights for three key categories: salary and wages, fringe benefits and supplies.[27] These three categories account for over 90 percent of direct cost expenditures, excluding other costs. The indirect cost category, while of significance for academic institutions, is ignored for purposes of the variance analysis because IMPAC files still contain data on that category. Thus, the variance of sample estimates of indirect cost weights need not be a criterion for choosing sample size and strategy since the weight for that category can be derived directly from the population data.

The confidence intervals reported in this section show, for example, that given a sample size and strategy, the data indicate there is a 95 percent probability that the sample mean will lie within an interval of plus or minus 3 percentage points around the true (population) mean.[28] A smaller confidence interval, of course, is better than a larger one. Given a target for the confidence interval, one can then make a choice with respect to the most suitable sampling strategy and sample size. Since each expenditure share (i.e. for salary and wages, fringe benefits and supplies) will have its own confidence interval, the ideal strategy will hopefully satisfy the target along more than one dimension.

Confidence intervals at the 95 percent level for simple random samples and stratified random samples of different sizes are shown in Tables 15(a) and 15(b). These tables only include the results for 1993. The findings for 1991 and 1995 are virtually identical and are presented in

---

[27] The variance estimates underlying the confidence intervals are presented in Appendix H.

[28] One can construct confidence intervals for any desired level of probability. Common alternatives to the 95 percent level of significance are the 90 percent and 99 percent levels of significance. The 95 percent confidence interval is

Appendix H. As shown in Table 15(a), the share of salary and wages in total direct costs, less other costs, was 0.6362 for the population of academic institutions in 1993. Repeated simple random samples of 10 academic institutions will lead to a number of estimates for the wage share and 95 percent of those estimates will lie within plus or minus 0.0408 points of the population share of 0.6362. In other words, repeat simple random sample of 10 will generate estimates such that 95 percent of them lie within the interval 0.5954 to 0.6770. The confidence interval gets smaller as sample size increases with the largest incremental benefit coming from an increase in the random sample from 10 institutions to 25 institutions.

Stratifying the sample of academic institutions offers clear benefits. A two strata sample of 10 institutions yields confidence intervals as low as those offered by a simple random sample of 25 institutions. However, there are very limited gains from expanding the number of strata from two to five. If a confidence interval of plus or minus 3 percentage points or better is adopted as the target for the sample of academic institutions, Table 15(a) shows that stratified samples of either 10 or 25 institutions would deliver very satisfactory results. A random sampling strategy would call for samples of 25 or more.

Turning to the universe of Top 100 academic institutions, Table 15(a) reveals that simple random samples of either 9 or 18 institutions lead to very narrow confidence intervals. Since the expenditure shares for the universe of the top 100 institutions are virtually the same as the shares for the overall population, small samples from the group of the top 100 can also be said to yield very satisfactory estimates of the population weights.

Findings for non-academic institutions are presented in Table 15(b). Since the population of non-academic institutions is both larger and more diverse, it is not surprising that the confidence intervals point to the need for larger samples of non-academic institutions. For instance, simple random samples of 10 from the population of non-academic institutions would yield wage share estimates with confidence intervals of 0.1133. That means the wage share estimate from the small random sample will fall anywhere in the range of 0.5289 to 0.7555 with

computed as Population Mean +/- 1.96*Standard Deviation, where Standard Deviation is the square root of the variance of the sample estimate.

Table 15(a)
Confidence Intervals at the 95 Percent Level of Significance for Selected Direct Cost Shares Academic Institutions, 1993

Confidence Intervals for the Population of Academic Institutions (N = 472)

		Random Samples					
Expenditure Category	*Expenditure Share*	n = 10	n = 25	n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6362	0.0408	0.0254	0.0174	0.0116	0.0088	0.0070
Fringe benefits	0.1554	0.0326	0.0203	0.0139	0.0093	0.0070	0.0056
Supplies	0.1401	0.0277	0.0172	0.0118	0.0079	0.0060	0.0048

		Stratified Samples: 2 Strata					
		n = 10	n = 25	n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6362	0.0258	0.0156	0.0102	0.0058	0.0033	0.0016
Fringe benefits	0.1554	0.0202	0.0122	0.0079	0.0045	0.0024	0.0011
Supplies	0.1401	0.0176	0.0107	0.0070	0.0040	0.0022	0.0011

		Stratified Samples: 5 Strata				
		n = 25	n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6362	0.0163	0.0100	0.0057	0.0030	0.0014
Fringe benefits	0.1554	0.0129	0.0079	0.0044	0.0023	0.0010
Supplies	0.1401	0.0111	0.0068	0.0039	0.0021	0.0010

Confidence Intervals for the Top 100 Academic Institutions (N = 100)

		Random Samples		
Expenditure Category	*Expenditure Share*	n = 9	n = 18	n = 27
Salary and wages	0.6360	0.0212	0.0143	0.0110
Fringe benefits	0.1569	0.0170	0.0114	0.0088
Supplies	0.1404	0.0144	0.0096	0.0074

Source: Joel Popkin and Company tabulations from IMPAC data.

Table 15(b)
Confidence Intervals at the 95 Percent Level of Significance for Selected Direct Cost Shares
Non-academic Institutions, 1993

Confidence Intervals for the Population of Non-academic Institutions (N = 1150)

Expenditure Category	*Expenditure Share*	*Random Samples* n = 10	n = 25	n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6422	0.1133	0.0712	0.0498	0.0344	0.0274	0.0231
Fringe benefits	0.1482	0.0761	0.0478	0.0335	0.0231	0.0184	0.0155
Supplies	0.1250	0.1241	0.0780	0.0545	0.0377	0.0300	0.0253
		Stratified Samples: 2 Strata					
		n = 10	n = 25	n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6422	0.1127	0.0708	0.0359	0.0209	0.0157	0.0118
Fringe benefits	0.1482	0.0537	0.0335	0.0174	0.0104	0.0077	0.0054
Supplies	0.1250	0.0738	0.0458	0.0243	0.0148	0.0108	0.0073
			Stratified Samples: 5 Strata				
			n = 25	n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6422		0.0443	0.0251	0.0158	0.0134	0.0096
Fringe benefits	0.1482		0.0221	0.0129	0.0084	0.0069	0.0052
Supplies	0.1250		0.0309	0.0188	0.0123	0.0100	0.0080
				Stratified Samples: 10 Strata			
				n = 50	n = 100	n = 150	n = 200
Salary and wages	0.6422			0.0256	0.0162	0.0121	0.0087
Fringe benefits	0.1482			0.0139	0.0088	0.0061	0.0043
Supplies	0.1250			0.0214	0.0136	0.0088	0.0060

Confidence Intervals for the Top 100 Non-academic Institutions (N = 100)

Expenditure Category	*Expenditure Share*	*Random Samples* n = 9	n = 18	n = 27
Salary and wages	0.6428	0.0435	0.0292	0.0225
Fringe benefits	0.1534	0.0296	0.0199	0.0153
Supplies	0.1361	0.0455	0.0305	0.0235

Source: Joel Popkin and Company tabulations from IMPAC data.

95 percent probability. It would take a random sample of 150 or more non-academic institutions to meet a target of plus or minus 3 percentage points for the expenditure categories in question. Even if the threshold is lowered to plus or minus 4 percentage points for non-academic institutions, a random sampling strategy would require a sample of 100.

As is the case with academic institutions, stratification of the sample also offers benefits over simple random sampling for non-academic institutions. However, it now takes five strata for the clearest benefits to emerge. A stratified sample of 25 to 50 would appear to provide the assumed degree of accuracy for non-academic institutions. There are virtually no gains from moving to a ten strata strategy in comparison to a five strata approach. Limiting the universe to the top 100 non-academic institutions suggests that a sample of 18 non-academic institutions might provide the optimum results.

One factor that offsets the apparent need for larger samples of non-academic institutions is their relatively small weight in the BRDPI. In BEA's calculations at the present time, non-academic institutions have an overall weight of 0.2107 in the BRDPI. In contrast, the weight for academic institutions is 0.5902.[29] Therefore, the impact of the non-academic component on the overall rate of change in the BRDPI is relatively minor and one may tolerate a higher degree of variance in the estimates of expenditure weights for the non-academic institutions. A higher tolerance for error would translate to smaller sample sizes. This issue is touched upon again in the following section in the context of simulations of the BRDPI using sample weights.

6.3. Estimates of the Extramural BRDPI Using Sample Weights

Another way to judge the effect of alternative sampling strategies is to re-estimate the BRDPI using sample weights and compare it to the BEA-estimated BRDPI. According to BEA, the extramural academic BRDPI increased 198.74 percent between 1979 and 1999 at an average annual rate of 5.62 percent. During the same time, the extramural non-academic BRDPI increased 170.53 percent at an average annual rate of 5.10 percent. If the sub-indexes for the BRDPI are aggregated without the use of expenditure weights, the 20-year change for the

[29] The remainder – 0.1992 – is accounted for by internal activities.

extramural academic BRDPI is only 174.48 percent at an average annual rate of 5.18 percent. The comparable figures for the extramural non-academic BRDPI are 164.78 percent and 4.99 percent per year respectively. Thus, expenditure weights do make a difference to the estimation of the BRDPI, especially its academic component, and depending on the sampling strategy used to estimate the weights there might be noticeable changes in the BRDPI.[30] The degree of the impact on the BRDPI could help choose among alternative strategies.

The main results are summarized in Tables 16(a) and 16(b) below. More detailed results are presented in Appendix H. Tables 16(a) and 16(b) show the 20-year and average annual rates of change in the extramural BRDPI using weights derived from the experimental samples discussed in Section 6.1 above. In principle, each one of those sampling strategies could have been applied repeatedly. For example, several 5 percent random samples of the population could have been taken and the BRDPI re-estimated for each to simulate the effect of repeated drawings of random samples of that size. However, that issue is effectively covered by the variance analysis discussed in the preceding section. The focus here is on a broad comparison across samples, with each sampling strategy being applied only once. Because the BRDPI is currently based on the 1993 IMPAC, the sample weights underlying the simulations of the BRDPI were also drawn from the 1993 IMPAC data.

The results are very encouraging. For the extramural academic BRDPI, the largest impact is registered in the case of non-probability samples restricted to the nine largest institutions (see Table 16(a)). However, even this effect is minor. Over the 1979-1999 period, weights derived from a sample of the nine largest academic institutions would have caused the average annual rate of increase in the extramural academic BRDPI to be 0.06 percentage points higher. Accumulated over the full twenty years, this would have meant an index level 3.66 percentage points higher than otherwise. In all but one other sample of academic institutions, the 20-year change is within 3 percentage points of the BEA estimate and the average annual rate of change is within 0.05 percentage points of the BEA estimate. Random samples of 9, 18 and 27 from the

[30] Expenditure weights would matter little if the sub-indexes in the BRDPI behaved the same over time. For instance, if all sub-indexes (indirect costs, wages and salaries, fringe benefits, etc.) moved at the same rate of 5 percent per year, the estimated change in the BRDPI will be 5 percent no matter which set of weights is used. That, however, is not the case with the BRDPI.

Table 16(a)
Cumulative and Average Annual Growth in the Extramural Academic BRDPI Under Alternative Sets of Sample Weights

Type of Sample	*Change in Extramural Academic BRDPI*		*Percentage Point Difference (BEA - Sample)*	
	Total: 1979-1999	Average Annual	Total: 1979-1999	Average Annual
Population weights (BEA)	198.74%	5.62%	---	---
Random samples:				
One in five (n = 88)	199.16%	5.63%	-0.41	-0.01
One in ten (n = 42)	197.48%	5.60%	1.27	0.02
One in twenty (n = 17)	198.19%	5.61%	0.55	0.01
Stratified samples:				
Two strata (n = 25)	202.03%	5.68%	-3.28	-0.06
Two strata (n = 51)	198.02%	5.61%	0.72	0.01
Two strata (n = 66)	199.19%	5.63%	-0.45	-0.01
Two strata (n = 107)	199.86%	5.64%	-1.11	-0.02
Two strata (n = 208)	199.81%	5.64%	-1.07	-0.02
Five strata (n = 28)	201.75%	5.68%	-3.00	-0.05
Five strata (n = 50)	198.16%	5.61%	0.58	0.01
Five strata (n = 82)	199.03%	5.63%	-0.29	-0.01
Five strata (n = 110)	199.77%	5.64%	-1.03	-0.02
Five strata (n = 216)	199.75%	5.64%	-1.01	-0.02
By Region	201.37%	5.67%	-2.62	-0.05
By Number of Awards:				
Top 9	201.69%	5.68%	-2.94	-0.05
Top 20	200.92%	5.66%	-2.18	-0.04
By Total Dollars Awarded:				
Top 9	202.40%	5.69%	-3.66	-0.06
Top 20	200.98%	5.66%	-2.24	-0.04
Top 100 Institutions:	200.27%	5.65%	-1.52	-0.03
Sample 1 (n = 9)	200.41%	5.65%	-1.67	-0.03
Sample 2 (n = 18)	199.95%	5.65%	-1.21	-0.02
Sample 3 (n = 27)	199.64%	5.64%	-0.90	-0.02

Note: The weights for all samples are drawn from the 1993 IMPAC files.
Source: Joel Popkin and Company.

Table 16(b)
Cumulative and Average Annual Growth in the Extramural Non-academic BRDPI Under Alternative Sets of Sample Weights

Type of Sample	*Change in Extramural Non-academic BRDPI*		*Percentage Point Difference (BEA - Sample)*	
	Total: 1979-1999	Average Annual	Total: 1979-1999	Average Annual
Population weights (BEA)	170.53%	5.10%	---	---
Random samples:				
One in five (n = 256)	169.00%	5.07%	1.53	0.03
One in ten (n = 125)	167.40%	5.04%	3.13	0.06
One in twenty (n = 74)	172.13%	5.13%	-1.60	-0.03
Stratified samples:				
Two strata (n = 26)	171.09%	5.11%	-0.56	-0.01
Two strata (n = 53)	169.15%	5.08%	1.37	0.03
Two strata (n = 75)	173.30%	5.16%	-2.78	-0.05
Two strata (n = 144)	170.67%	5.10%	-0.14	0.00
Two strata (n = 277)	170.89%	5.11%	-0.36	-0.01
Five strata (n = 27)	171.96%	5.13%	-1.43	-0.03
Five strata (n = 49)	170.58%	5.10%	-0.05	0.00
Five strata (n = 109)	170.95%	5.11%	-0.43	-0.01
Five strata (n = 116)	170.71%	5.11%	-0.18	0.00
Five strata (n = 206)	171.68%	5.12%	-1.15	-0.02
Ten strata (n = 52)	170.80%	5.11%	-0.27	-0.01
Ten strata (n = 145)	173.00%	5.15%	-2.47	-0.05
By Region	170.02%	5.09%	0.51	0.01
By Number of Awards:				
Top 9	170.63%	5.10%	-0.10	0.00
Top 20	170.88%	5.11%	-0.35	-0.01
By Total Dollars Awarded:				
Top 9	170.63%	5.10%	-0.10	0.00
Top 20	170.14%	5.09%	0.38	0.01
Top 100 Institutions:	171.44%	5.12%	-0.91	-0.02
Sample 1 (n = 9)	172.84%	5.15%	-2.31	-0.04
Sample 2 (n = 18)	177.11%	5.23%	-6.58	-0.13
Sample 3 (n = 27)	169.84%	5.09%	0.69	0.01

Note: The weights for all samples are drawn from the 1993 IMPAC files.
Source: Joel Popkin and Company.

top 100 academic institutions perform very well, delivering a 20-year change within 1.67 percentage points of the BEA and an annual average rate within 0.03 percentage points. These results outperform those from non-probability samples, such as the top 9 institutions, and are just as good as those obtained from larger stratified samples.

The results for non-academic institutions are similar in the sense that most samples perform very well (see Table 16(b)). As is the case with academic institutions, only two samples lead to BRDPI estimates whose change over the 20-year period differs from the official estimate by over 3 percentage points. Those are the simple one-in-ten random samples and the random sample of 18 from the top 100 non-academic institutions. In all other samples for non-academic institutions, the average annual rate of change is within 0.05 percentage points of the BEA estimate.

In sum, almost all sampling strategies yield BRDPI estimates that are very close to the original. A corollary is that the results across the different samples are very similar and there is little gain from larger sample sizes, at least as far the BRDPI estimates are concerned. However, it is important to keep in mind that the point estimates of the weights that underlie the simulations in Tables 16(a) and 16(b) are subject to sampling variance and additional drawings of the same samples could have led to different results. Unfortunately, attempting to simulate the BRDPI with repeat drawings of a large number of sampling strategies would be an arduous task. It was decided instead to try two experimental simulations based on some assumptions regarding the outer limits of change in expenditure weights.

The first experiment is based on the confidence intervals for sample means estimated from simple random samples of 10 institutions. It was assumed that the sample weights for salary and wages and fringe benefits lie at the outer limits of their confidence intervals. For academic institutions, this meant adding (or subtracting) 0.045 points to (from) the population salary and wages share and 0.035 points to (from) the fringe benefits share. For non-academic institutions, this meant adding (or subtracting) 0.120 and 0.075 points to (from) the salary and wages share and fringe benefits share respectively. The weights for these two categories were either increased simultaneously or decreased simultaneously. Weights for all other direct cost elements were reduced (or increased) in proportion. For academic institutions, the indirect cost share was left unchanged. The resulting academic and non-academic extramural BRDPI indexes were then combined with the internal activities BRDPI using BEA weights. That leads to an estimate of the impact on the total BRDPI in the event of large differences between sample estimates and populations weights for salary and wages and fringe benefits. The results are shown in Table 17.

Table 17 below shows the 20-year and average annual rates of change in the extramural academic and non-academic BRPDI, the internal activities BRDPI and the total BRDPI derived from official and simulated estimates of the indexes.[31] BEA estimates show that the total BRDPI increased 154.34 percent between 1979 and 1999 at a rate of 4.78 percent per year. If the shares for salary and wages and fringe benefits are increased to their outer limits as defined above, the 20-year change in the total BRDPI is now estimated to be 162.30 percent at an average annual rate of 4.94 percent. Conversely, if the expenditure weights for these two categories are reduced to the lower ends of their confidence intervals, the total BRDPI is estimated to have increased 146.73 percent between 1979 and 1999 at an average rate of 4.62 percent per year. Thus, on an annual basis, even this "worst-case" scenario makes for a difference of only plus or minus 0.16 percentage points in the total BRDPI. For academic institutions alone, the BRDPI growth rate is changed by plus or minus 0.19 percentage points on an annual basis and by a total of 10.5 percentage points over the 20-year period. For non-academic institutions alone, the effect is plus or minus 0.33 percentage points on an annual basis or 15.5 percentage points combined between 1979 and 1999. These results are encouraging in the sense that the tolerance for errors in expenditure share estimates, especially for non-academic institutions, appears to be quite high. Even large swings in the weights for two important expenditure categories produce modest changes in the total BRDPI.

The second experiment that was tried was to ascertain the extent of change in the weights for salary and wages and fringe benefits that is needed to generate an absolute change of 0.25 percentage points in the annual rate of growth in the total BRDPI. It was decided to keep the change in weights for non-academic institutions at the same level as described for the first experiment above. The weights for academic institutions were then altered from the population

[31] See Appendix H for more detailed results.

Table 17
Cumulative and Average Annual Growth in the BRDPI under Assumed Deviations of Weights for Wages and Fringe Benefits from Population Weights

	Change in BRDPI	
	Total: 1979-1999	Average Annual
Internal Activities BRDPI:	66.57%	2.58%
Extramural Academic BRDPI:		
Population weights	198.74%	5.62%
"Maximum" weights	209.28%	5.81%
"Minimum" weights	188.75%	5.45%
Extramural Non-academic BRDPI:		
Population weights	170.53%	5.10%
"Maximum" weights	188.06%	5.43%
"Minimum" weights	154.54%	4.78%
Total BRDPI:		
Population weights	154.34%	4.78%
"Maximum" weights	162.30%	4.94%
"Minimum" weights	146.73%	4.62%

Note: "Maximum" weights refers to the increase in the population weights for salary and wages and fringe benefits to the upper end of the 95 percent confidence interval as determined by a random sample of 10 institutions. "Minimum" weights refers to the decrease in the population weights for salary and wages and fringe benefits to the lower end of the 95 percent confidence interval as determined by a random sample of 10 institutions. See the text and Appendix H for additional details.

Source: Joel Popkin and Company tabulations based on IMPAC data and BEA price index data.

means until the desired change in the total BRDPI was obtained. The required changes in the expenditure weights for academic institutions proved to be as follows: plus or minus 0.08 percentage points for the salary and wages share and plus or minus 0.065 percentage points for the fringe benefits share. As before, the two weights were increased or decreased simultaneously with offsetting proportional changes in the weights for other direct cost categories. These changes in the weights are very large and are nearly double the confidence intervals from the random sample of 10 academic institutions. Under these assumed changes the annual growth rate in the extramural academic BRDPI is affected by 0.33 percentage points. The average annual growth rate in the total BRDPI is changed by 0.24 percentage points and the 20-year impact on the total BRDPI is 12 percentage points.

In sum, the various simulations show that even large sampling errors in the estimation of expenditure weights for the BRDPI are unlikely to have a significant effect on the rate of change in the BRDPI. Because of the randomness in the sampling error, a strategy that causes the BRDPI to increase faster in its first application could very well cause the BRDPI to increase slower in the next go around. In other words, it is unlikely that a chosen sampling strategy will forever cause the BRDPI to increase slower or faster.[32] In the long run, the average deviation from the "true" rate of change in the BRDPI should be near zero under any sampling strategy. The choice among the different strategies depends on the extent of the short-run fluctuations around the true trend line that is considered permissible. In the context of the analysis presented in Table 16, it may be decided that sample weights should yield a BRDPI with a 20-year change within 5 percentage points, or an average annual rate of change within 0.09 percentage points, of the "true" BRDPI. By that criterion, virtually any of the sampling strategies tried here pass the test and one can choose one that is the simplest and cheapest to administer.[33]

[32] That is the key difference between the issues at hand here and the debate over the bias in the CPI. The bias in the CPI is said to be persistently unidirectional and caused by factors such as substitution bias, changes in the quality of goods and services, introduction of new goods, etc. Even if all sources of bias were removed from the CPI it, like the BRDPI, will still be subject to sampling errors resulting from the measurement of expenditure weights and prices. Schultze and Mackie (2002) is an excellent reference on the issues surrounding the CPI.

[33] It is worth noting that while sampling error may not cause much of a change in the BRDPI as a switch is made from population weights to sample weights, a bigger change could result from the fact that the IMPAC data capture planned expenditures and survey data capture actual expenditures.

7. Options for the BRDPI Expenditure Survey and their Estimated Cost

This paper began with a listing of the major issues that needed to be resolved before a design for the BRDPI expenditure survey could be finalized. The analyses of the IMPAC data and the test survey have helped settle all of those issues. The unit of observation for the survey must be the institution. Using the award as the unit of observation is ruled out due to the burden it would impose and the unavailability of data on individual awards from many institutions. The expenditure survey cannot be counted upon to expand the number of categories for which expenditure data can be collected over and above the number currently contained in the BRDPI. If desired, a limited amount of information may be collected on secondary issues, such as, the proportion of salary and wages in an institution subject to the NIH salary cap. The test survey proved successful in its ability to draw complete and accurate answers from the respondent institutions. It is recommended that future versions of the BRDPI expenditure survey adhere closely to the questionnaire and format of the test survey as shown in Appendix I. Of course, some of the questions included in the test survey lose some of their relevance for a final survey and may be modified or omitted as deemed suitable.

The remaining decisions regarding the BRDPI expenditure survey concern the following issues: (a) the choice of a sampling strategy or strategies; (b) the frequency and timing of the survey; (c) respondent burden and OMB clearance; (d) the implications of the expenditure survey for computation of the BRDPI by the BEA; (e) the estimated cost of the final survey. The resolution of these issues is discussed in turn below.

7.1. The Choice of Sampling Strategy

Estimates of sampling error and simulations of the BRDPI with sample weights showed conclusively that fairly small sample sizes can fulfill the goals of the expenditure survey. In particular, no more than 9 to 18 academic institutions and 9 to 27 non-academic institutions are needed for the sample. A larger sample of non-academic institutions is potentially needed for two reasons. One reason is the population of non-academic institutions is both bigger and more diverse than the population of academic institutions. The second reason is the higher sampling error of estimates for non-academic institutions. On the other hand, non-academic institutions control only about 20 percent of total funding and, therefore, have very limited impact on the BRDPI. That means that one could tolerate a higher level of sampling error and a smaller sample size for non-academic institutions. Thus, the recommended sample of non-academic institutions could also lie in the range of 9 to 18 institutions.

Because of the skewed nature of NIH funding, the focus of the sample of both academic and non-academic institutions should be on large institutions. That is the most efficient way of covering the largest possible share of extramural funding provided by NIH. For instance, the nine largest academic institutions alone control 25 percent of the total funding given to academic institutions. One possibility is to stratify the sample by size of institution, where size is determined by the total number of dollars awarded to the institution by NIH. No more than two strata are needed for academic institutions but the non-academic institutions should be grouped into five or more strata. There are two alternatives to stratification. One is to use non-probability sampling and select the desired number of the largest institutions. This could mean sampling only the top 9 or top 20 institutions based on total dollars awarded. The other alternative is take simple random samples from a population defined to consist only of the top 100 institutions. The option of sampling from the top 100 institutions offers a very simple, efficient and accurate method for estimating expenditure weights for the BRDPI. It also offers the advantage of distributing the burden of the survey across different institutions from one time period to the next, a feature lacking from the non-probability samples.

Small sample sizes do increase the importance of response rate. Fortunately, a small sample also makes it feasible to generate a high response rate because personal contact and follow up is easier and less expensive with a small group. The test survey resulted in seven responses from nine mailings. With adequate follow up, there is no reason to suspect that the response rate for the final survey will be any less.

7.2. The Frequency and Timing of the Expenditure Survey

The frequency of the survey depends in part upon the need to update the expenditure weights for the BRDPI. As shown by the examination of the IMPAC data, BRDPI expenditure weights barely changed between 1991 and 1995. Also, there is no intention at the present time to change the formula of the BRDPI from a Laspeyres price index to a Paasche index or another index that requires current period expenditure weights for its computation.[34] Thus, updating the BRDPI expenditure weights about every five years ought to be sufficient.

To avoid the need for OMB clearance, the expenditure survey could be conducted over a period of three to five years. The sample size in any single year could be restricted to nine institutions. One potential design is as follows: Survey nine academic institutions in the first year and, if necessary, follow up with a survey of nine more in the second year. The same tactic could be repeated with non-academic institutions over the next two years. A fifth year could be reserved to make up for non-response or take account of unforeseen complications. Under this plan, each group of institutions – academic and non-academic – will be surveyed every five years albeit not at the same time.

Because fiscal years for most institutions end in the months of June through September the expenditure survey should be timed for field distribution in the winter months, perhaps January or February. The accounting offices at the institutions are likely to be less busy during these months than during the beginning and end of fiscal years.

7.3. Respondent Burden and OMB Clearance

OMB clearance will not be required if the survey is stretched over a period of time as discussed above. Respondent burden is measured by the number of hours it would take a respondent to complete the BRDPI expenditure survey. The test survey indicates that the average burden could be up to nine hours per respondent. Based on this figure a survey of nine

[34] Other examples of price indexes that require current period weights are the Fisher index and the Tornqvist index. Chain-weighted indexes also require constant updating of weights.

institutions implies a total burden of 81 hours and four-year cycle covering 36 institutions would lead to a total burden of 324 hours. The respondent burden will also depend upon whether or not questions about the salary cap are included in the final survey. Some respondents to the test survey commented that they spent much of their time gathering the data needed to answer the salary cap question. Without this question the burden for may fall closer to five hours than nine hours. Thus, the estimates of respondent burden given above should be regarded as the upper limits.

7.4. Implications of the Expenditure Survey for the BEA Methodology

The small sample sizes will require the BEA to make changes in its methodology. Data on the expenditure distribution within individual institutions will no longer be at BEA's disposal. Those data are used by BEA to compute the wage, fringe benefits, indirect cost and MDTC indexes for academic institutions. The only thing BEA will be able to continue doing as before is the aggregation of the indirect cost index across institutions. That is because the IMPAC files still contain information on dollars awarded to individual institutions to cover indirect costs. But new methods will have to be used to compute the wage, fringe benefits and MDTC indexes for academic institutions. The analysis showed that several reasonable alternatives are available to BEA and none of them will have an adverse effect on estimates of the BRDPI.

It was shown that, instead of relying on data on wages and fringe benefits, BEA can use the across-institution distribution of total dollars awarded or total direct costs to estimate the wage and fringe benefits indexes. The across-institution distribution of total dollars awarded is nearly perfectly correlated with the across-institution distribution of wage and fringe benefits dollars.[35] Data on total dollars awarded and total direct costs are still available in the IMPAC files. For the MDTC index, BEA can substitute estimates of population weights for institution weights to aggregate the MDTC component indexes. This method works remarkably well for the large majority of institutions. A reason for that is the BEA sample of institutions consists almost entirely of large institutions. As shown by the research, that is a relatively homogenous group of

[35] Also, as discussed in Section 5, there isn't much across-institution variance in the price indexes BEA computes. This means that their aggregation is not too sensitive to the choice of weights.

institutions and the expenditure distribution within each of these large institutions does not differ too much from the expenditure distribution for the overall population.

There is one aspect of its methodology that BEA may have to sacrifice regardless of sample size. BEA collects wage data separately for medical and non-medical research personnel. Within institutions with both medical and non-medical faculty, the indexes are aggregated using the proportions of research done by the two types of personnel. These weights are estimated from the IMPAC files. However, the BRDPI expenditure survey does not intend to include a question on this particular division of research expenditures. BEA may have to abandon this approach or, as a proxy, use weights based on the medical/non-medical distribution of total direct costs. That information is still available in the IMPAC files.

7.5. Estimated Cost of the BRDPI Expenditure Survey

The total cost of the final BRDPI expenditure survey is expected to be modest. The main reason, of course, is that the sample size, for either academic or non-academic institutions, will be small. The cost estimates presented below are based on tasks most directly associated with the conduct of the survey. Those include developing a mailing list, printing and distributing the survey, follow-up to the survey to increase response rate, data entry and editing, tabulating the data, and miscellaneous supplies and labor. OMB clearance, in the event it is required, is not budgeted for within the estimates below. Because of the modest size of the sample, the survey is assumed to be produced "in-house" as opposed to being contracted to a printing firm. Another assumption is that the final survey will be similar to the test survey so that substantial amount of work is not required to develop the questionnaire. Statistical analysis of the data, such as estimating the variance of the estimates, is also outside the scope of the budget estimates that follow. It is possible that the statement of work that may be written for the final expenditure survey may include that task among others. However, it is not possible to price those task elements without knowledge of the final statement of work. For these reasons, the cost estimates below should be treated as approximate and subject to change as the final tasks are clarified.

Table 18 below presents cost estimates for the BRDPI expenditure survey assuming sample sizes of 9, 18 and 27.[36] It is assumed that a sample of 18 means two surveys of nine institutions spread over two years. Similarly, the sample of 27 institutions consists of three surveys of nine institutions each. A consequence of this structure is that there are no economies of scale for many tasks. For instance, it is assumed that a sample of 27 requires the production of three reports, one at the end of each of three surveys.

The major tasks prior to the conduct of the survey are finalizing the design of the survey, drawing the sample from the IMPAC files, and developing the mailing list. Finalizing the design of the survey refers to putting the finishing touches on the formatting of a survey instrument that is not too different from the test survey. Alternatively, this could involve the redesign of the survey for distribution via e-mail. This cost element does not vary with sample size. Developing the mailing list is a labor-intensive task that includes the verification of the address list so that mail (or e-mail) is not misdirected. However, its cost is limited by the modest size of the samples.

Table 18
Estimated Cost of the BRDPI Expenditure Survey

Cost Element	*Sample Size*		
	9	18	27
Finalize survey design	$2,310	$2,310	$2,310
Draw sample of institutions	$770	$770	$770
Develop mailing list	$165	$330	$495
Printing and mailing	$73	$146	$219
Follow-up to survey	$55	$110	$165
Data entry and editing	$165	$330	$495
Data tabulations	$825	$1,650	$2,475
Report writing	$2,310	$4,620	$6,930
Miscellaneous tasks	$2,310	$3,465	$4,620
Total	$8,983	$13,731	$18,479

[36] The cost estimates assume fully loaded hourly rates of $110/hour for the project director, $55/hour for a research analyst, and $40/hour for other labor.

Survey printing and mailing costs are also modest. This cost component would diminish to virtually zero if the survey is distributed via e-mail instead of by the post. The follow-up to the survey involves contacting non-respondents and/or answering questions regarding the survey. It is assumed that this is done either on the phone or by e-mail. Data entry and editing is not onerous due to the small sample size and the limited extent of the questionnaire. For the same reasons, not many tabulations could be derived from the survey data and are presumed to consist mostly of information on expenditure weights. It is possible to supplement the expenditure survey data with analysis from the IMPAC files regarding the characteristics of the institutions included in the survey and the awards they receive. However, the cost of this analysis, if called for, is not included in Table 18.

Report writing consists of the presentation of the results and general descriptions of the survey procedures, response rates, etc. Again, these costs could rise if supplemental analysis from the IMPAC files must be conducted and reported. The final line item represents an allowance for miscellaneous or unanticipated tasks. In total, the cost of the survey is expected to range from approximately $9,000 to $18,500 depending on the choice of sample size. Some of the cost elements, in particular the survey design, represent "one-time" costs and will not be incurred with regularity in the future. Repeat surveys may also cost less because many of the larger institutions are likely to make repeat appearances in the sample and that will serve to reduce the burden of developing a mailing list. Similarly, data editing and tabulating will become more of a routine over time. However, as indicated in the preceding discussion, survey costs could also increase substantially if supplemental analysis from the IMPAC files is called for and/or the survey data are subjected to exhaustive statistical analysis.

Bibliography

Bureau of Economic Analysis, Government Division, "Biomedical Research and Development Price Index," Draft, December 1999.

Bureau of Economic Analysis, Government Division, "Biomedical Research and Development Price Index for Academic Performers," November 1983.

Bureau of Economic Analysis, Government Division, "Analysis of Overhead Costs for Biomedical Research and Development," April 1980.

Bureau of Economic Analysis, Government Division, "Evaluation of the NIH Biomedical Research and Development Price Indexes," October 1979.

Kish, Leslie, Survey Sampling, John Wiley and Sons, New York: 1965.

Schultze, Charles and Mackie, Christopher (eds.), At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes, National Academy Press, Washington, DC, 2002.

The College Board, College Handbook, 2001, The College Board, New York.

The Princeton Review, Complete Book of Medical Schools, 2001, Random House, Inc., New York.

The National Institutes of Health, Office of Science Policy and Technology Transfer, Office of the Director, "Biomedical Research and Development Price Index," June 1993.

The National Institutes of Health, Office of Science Policy and Legislation, Division of Planning and Evaluation, Economic and Resource Studies Branch, "Comparison of Award versus Expenditure for Direct Costs: A Sample of 840 R01 Grants from Eight Institutions for Fiscal Year 1985," September 1989.

Yamane, Taro, Elementary Sampling Theory, Prentice-Hall, Englewood Cliffs, NJ: 1967.